

# How to make the NPT more effective

**The defection of North Korea from membership of the Non-Proliferation Treaty is a conundrum for the review conference two years from now. Built-in sanctions are now required, as are ways of enforcing membership.**

WILL North Korea be the next nuclear power to declare itself as such? That prospect is easily inferred from its government's decision to withdraw from the Nuclear Non-Proliferation Treaty (NPT). By all accounts, North Korea has found irksome the inquisitiveness of the inspectors looking into the functions of its declared nuclear facilities and their assertion of their right (under the new guidelines of the International Atomic Energy Agency, or IAEA) to inspect facilities that may be relevant, but which have not been so declared. Meanwhile, there are unconfirmed reports that North Korea has enough plutonium to make a bomb.

Caution is, for the time being, the best option. Although North Korea's interest in nuclear energy has been openly advertised for many years, the assumption that its withdrawal from the NPT implies that it is about to become a military nuclear power is not proven. Since the end of the Korean war in the early 1950s, North Korea has been thoroughly unpopular in the West; it is still separated from the rest of the Korean Peninsula by a demilitarized zone declared as part of the armistice. The United States has pointedly refrained from appointing an ambassador. Would such an impoverished country seek further ostracization?

But this, unfortunately for us all, is a field in which facts are less important than perceptions. Whatever the truth about North Korea's bomb-making capacity, its neighbours' guesses as to what may be afoot there will determine their responses. South Korea, the most immediately interested, will no doubt wish to keep its options open. And while Japan, well within the range of the missiles North Korea is known to have, will not lightly abandon its traditional policies of being free from nuclear weapons, North Korea's abandonment of the NPT cannot but raise awkward questions there. Even China, a much more powerful nuclear power, cannot be unconcerned; nuclear superpowers are traditionally anxious that lesser fry do not drag them into trouble regionally.

The best hope for an immediate resolution of the problem of North Korea is that China should be persuaded to intervene. The days have passed, it seems, when China argued publicly that nuclear stability required that all states should have bombs. But that can be only a stop-gap solution. North Korea's defection from the NPT raises broader questions, all of which are relevant to the conference of the parties to the NPT that must be held within just over two years from now, at which the member states must agree that the treaty should continue to bind them, and in what manner.

To let the NPT lapse would not necessarily be disastrous,

but would saddle such superpowers as survive into the second half of the decade with regional responsibilities that they would imperfectly discharge. (Just think of what would happen in the Middle East.) But strengthening the safeguards regime, necessary though that may be in some respects, would be costly and would not suffice to halt the further spread of nuclear weapons as the case of Iraq, still a member of the NPT, amply demonstrates. Changes in the obligations required of members of the NPT are also needed.

Extinguishing the right of any state to withdraw from the NPT is an obvious amendment. In principle, the treaty offers non-nuclear states technical assistance with the development of civil nuclear power, whose benefits may be enduring and which may also help with the development of weapons. Why should such a power be free to enjoy the benefits and then walk out when the obligations become onerous? To give that argument more force, the IAEA should be asked to dust off its prospectus for technical assistance. With anxiety about global warming in the air, it will not be long before the economic interest of nuclear power is revived again.

The time has also come when breaches of the NPT require built-in sanctions. Those now applicable to Iraq are defined by resolutions of the United Nations Security Council. In principle, there is no reason why similar procedures should not be followed with North Korea which, if the darkest inferences are valid, must have been in breach of the NPT during its brief membership of the treaty. Everything hangs on whether the international community is serious in its protestations that the spread of nuclear weapons is potentially dangerous. If so, it could even go a significant step further, requiring NPT membership of states that do not now belong, India and Pakistan, for example, and even Israel. That, of course, would be a radical departure from past practice, but not one incommensurate with the importance of what should be the international community's objectives. □

## Back-door regulation

**The British government seems to be hankering after direct control of the universities whose costs it meets.**

BRITISH governments are notoriously able to manipulate their parliament's procedures, but they can rarely have done so as meanly as in their attempt, this month, to take powers to direct the behaviour of British universities and



the funding councils that now support them. The government has smuggled into the final version of a bill ostensibly concerned with the management of schools opting for central rather than local government control a provision that will allow central direction of all educational institutions supported by the central government, universities included. This broad form of the provision was not discussed during the committee stage of the bill in the House of Commons.

The issue has long and contentious roots. During the long debate on the 1988 Education Act, which created (among a multitude of other things) two funding councils for universities and polytechnics respectively (since merged and then split into three, for England, Scotland and Wales separately), the government fought hard for powers to direct the policies of the funding councils and the institutions dependent on them. Last year, the same fight was fought during the Further and Higher Education Bill's proceedings. Hitherto, the government has been denied its powers. So now, as unostentatiously as possible, it is trying again.

Given parliamentary procedure, only the House of Lords can now deny the government on this third occasion. It will have had its first chance on Tuesday this week, when the principles of the House of Commons's version of the bill will be considered. But the mere circumstance that a fight over the right of academic institutions to run themselves as they choose should have become an annual event is depressing, to say the least of it.

But is it not right that the government, which supports most British institutions of higher education, should have powers to regulate their conduct of affairs? That is what the government will say. Up to a point, that is of course correct. But the legislation of the past few years already ensures that universities are sufficiently accountable to their immediate paymasters, the funding councils. Formally, they are free to pursue idiosyncratic and even eccentric policies, but only at their own peril — that of jeopardizing their survival. What other straitjackets does the government have in mind?

Only a little imagination suggests lurid possibilities. Having willed (last year) that the number of British universities might be doubled, the government may now take it into its head that mergers between institutions would be tidy. Or it might require that universities should follow other public institutions in the fashion of 'market-testing' (which means seeking external contractors for some of the work it does) about to be decreed for other parts of the public sector should be followed in universities as well. Or does Mr John Patten, the Secretary of State for Education, have it in mind that universities should be required to publish indicators of their own performance in the education of the young?

Patten may of course have different things on his mind. The plain truth is that it is improper that he should be given the powers he has now again sought without saying what his intentions are. Until he does so, and his intentions are properly discussed, he will not be able to slough off the

charge that he is a centralizing minister, prepared in the last resort to substitute his own judgement for that of those who now function thanklessly as the trustees of universities and the mechanisms for subsidizing them. □

## Hacking in court

**Is computer hacking criminal or merely another endearing kind of mischief?**

"AN unlikely villain" is how one London newspaper described Paul Bedworth, the nineteen-year old computer buff cleared last week by a British court of several counts of conspiracy after hacking his way into hundreds of computer systems worldwide. The newspaper saw fit to add that the student, now at the University of Edinburgh, who cost the *Financial Times* more than £20,000 in telephone bills (hackers require onward telephone connections from their immediate victims) and who spent hours on end poring over the files of Belgian cancer patients, "has never had a girlfriend". It reported that he would like one, but "would not know" how to go about it. Applications on a postcard, please, and to the newspaper.

Bedworth's acquittal is no surprise, much as it has offended lawyers and those who lobbied for the 1990 Computer Misuse Act. Regardless of Bedworth's lawyers' contention that his enthusiasm for computers constituted an addiction, society — and perhaps British society particularly, whose affection for the maladjusted myopic figure of the scientist with the mischievous streak has not waned since the inventors of the bouncing bomb and the Enigma decoder helped it onto the winning side in the Second World War — has always had a soft spot for hackers.

Only in 1990 did hacking become a criminal offence. Since then, there has been a handful of prosecutions, the huge cost to British businesses notwithstanding. Alistair Kelman, Bedworth's defence counsel, agrees that British society regards computer hacking as endearing mischief-making. "No jury would equate the activities of this schoolboy [as he was] with those of ram-raiders or joy-riders", he says.

The public perception has always been that the damage inflicted by these enthusiasts, with their deceptively neanderthal postures and obsessions with the literary works of J. R. R. Tolkien, amounts to not much more than the random inconvenience of having to sit next to one of them in a train. And perhaps that is all computers deserve for turning modern life into the mosaic of bar codes and bank statements it has become.

The truth is not that Bedworth is an unlikely villain, but that computer systems and the companies that install them have come to seem deserving victims in the estimation of the modern media. Provided that hackers stop short of defrauding pensioners of their life savings, they can expect an easy ride not only from British newspapers but from the courts. The difficulty is that the Bedworth case is not relevant to the serious hacking that goes on, where people are paid for what they do and are smart enough not to be found out. □

# MicroGeneSys drops out of NIH trial for AIDS vaccine

**Washington.** The AIDS vaccine that the US Army intends to use exclusively in a \$20-million trial programme has been withdrawn by its maker from the main comparative trials being organized by US public

health agencies in a move that deepens already bitter divisions within the AIDS research community.

MicroGeneSys of Meriden, Connecticut, earlier this month withdrew its gp160 vaccine from trials planned by the AIDS Clinical Trials Group

**Franklin Volvovitz**

(ACTG) of the US National Institutes of Health (NIH). The company and its development partner, Wyeth-Ayerst Research, objected to the proposed dosing schedule, the duration of the study and the criteria being used to evaluate the vaccine and said that the withdrawal was "driven primarily by the needs of our development program".

The gp160 vaccine, known commercially as VaxSyn, consists of the outer envelope protein of the AIDS virus and is intended to boost the immune systems of people already infected with HIV.

Although NIH has declined to wage a public war of words against MicroGeneSys, officials are privately furious at the company's withdrawal and do not accept its explanation. "This was the only study giving direct comparability between rival vaccines", says one. "But put yourself in their shoes: why would they want to help generate data showing that their vaccine offers less immunogenicity than others?"

The president of MicroGeneSys, Franklin Volvovitz, says that the suggestion that his company fears the results of the proposed trials is "preposterous". "It's bizarre for them to say these things", says Volvovitz about NIH. "We disagreed with the primary and secondary end-points set for protocol 214. As the firm with the largest database, our views should be considered."

Volvovitz says that MicroGeneSys has been repeatedly excluded from another NIH competitive study, protocol 209, that monitors the effect of vaccines on patients' response to the AIDS drug AZT. Last

summer, the team running the protocol told the company that its limited resources would be best used to study a single vaccine developed by Genentech.

MicroGeneSys's successful effort last autumn to persuade Congress to give the US Army \$20 million for large-scale trials of its gp160 vaccine has aroused anger and jealousy within the AIDS research community (see *Nature* 360, 94; 1992). Critics of the special appropriation, which drew energetic support from AIDS activists, believed that clinical trials were premature.

That assessment was endorsed in November by a panel of experts convened by the NIH director, Bernadine Healy, but the panel nonetheless agreed that the Army programme should proceed "on compassionate grounds" as long as more than one vaccine is used. But the Army, which must notify Congress by 6 April of its plans, is now expected to propose stipulations for inclusion in its clinical trials that only the MicroGeneSys vaccine can meet.

In a further development, Colonel Robert Redfield, an Army scientist closely involved in the design of the test programme, is said to have been cleared by an internal Army investigation into allegations that he had overstated the effectiveness of the gp160 vaccine at a public meeting in Amsterdam last July. Redfield declined to comment, saying that "I'm trying to stay out of the gossip pages and do my science".

NIH is now likely to proceed with a protocol 214 based on comparative trials between gp120 vaccines produced by two California companies, Genentech and Chiron.

**Colin Macilwain**

## Clinton accepts court's ruling on Antarctica

**Washington.** The US government has abandoned its longstanding opposition to applying a 23-year-old environmental law to activities in Antarctica, a reversal that the National Science Foundation (NSF) says will require it to spend hundreds of thousands of dollars a year on unproductive administration instead of on science.

Last week, President Bill Clinton rejected a plea by several federal agencies to contest a recent ruling by the US Court of Appeals involving the construction of an incinerator at a US base at McMurdo Station (see *Nature* 361, 482; 1993). The court had ruled on 29 January that NSF must follow the requirements of the National Environmental Protection Act (NEPA). A lower court will now decide whether NSF has acted properly in building the incinerator.

The government has long claimed an exemption from NEPA for reasons of national security and for matters of state covering such activities as overseas military bases and economic development projects in the developing world. In deciding not to file an appeal, Clinton struck a compromise by accepting NEPA's jurisdiction over Antarctica while insisting that "the adminis-

tration does not embrace language in the opinion that may be interpreted to extend beyond this holding".

The Environmental Defense Fund, which brought the suit in 1991, says that there is no reason for NSF to be exempt from a law that it must follow in the United States. But NSF officials see the decision as a 'litmus test' that allows the president to flaunt his environmental convictions at the expense of a science agency that insists it already follows procedures at least as strict as those under NEPA. To emphasize that point, the National Science Board, NSF's governing body, last week passed a resolution affirming its support for a 1991 agreement intended to protect the Antarctic continent.

NSF estimates that it will have to employ three or four people to conduct and write up the environmental impact statements required by NEPA. The act also makes NSF vulnerable to what its director, Walter Massey, called "frivolous suits" by environmental groups that could impede research. "We do not feel that we will have to change the way we operate in Antarctica", he adds, "but the extensive documentation will cost a lot of money for a little agency like NSF."

**Jeffrey Mervis**

## Geneticists win Crafoord Prize

Two scientists who have developed radically different approaches to the genetic basis of behaviour in animals have been jointly awarded this year's Crafoord Prize by the Royal Swedish Academy of Sciences. The SKr2.6 million (US\$300,000) prize is to be shared by Seymour Benzer of the California Institute of Technology and Richard Hamilton of the University of Oxford. Benzer is best known for his studies of the genetic and neurophysiological basis of behavioural mutants in the fruit fly *Drosophila melanogaster*, and Hamilton for his theoretical work on the way in which kin selection and genetic relationships act as prerequisites for the evolution of altruistic behaviour.

**David Dickson**



# Curtain falls on Britain's nuclear structure facility

**London.** A meeting earlier this month at Daresbury in Cheshire to celebrate the achievements of the Nuclear Structure Facility (NSF) was also an occasion to mourn its premature death. On Monday (29 March), the last ions will be accelerated down the 70-metre high Van de Graaff accelerator to smash into nuclei at the base of the NSF, which fell victim to a financial crisis two

the ion-beam crystallography work. This so-called medium-energy ion-scattering facility (MEIS) will carry out structural investigations of the surfaces and sub-surface regions of materials.

Daresbury officials claim that MEIS will be "equal to the best facilities elsewhere in the world". But physicists working on the NSF point out that its predecessor has been a world leader. "The 'me-too' argument seems somewhat feeble", says John Sharpey-Shafer, who believes that closing NSF was "scandalous".

Even more jobs would have been lost if the NSF's demise had not coincided with rationalization at the nearby ICI Materials Research Centre. As part of this process, ICI has agreed to transfer to Daresbury a high-resolution X-ray photoelectron spectrometer, which carries out studies of the surface of materials at the molecular level.

The machine, unique in Britain, will become the focus of a new group known as the Research Unit in Surfaces, Transforms and

Interfaces (RUSTI). About ten scientists and technical staff from the NSF will be transferred to the new unit; one-third of the available time will be used by ICI and two-thirds will go to SERC grant-holders and to British companies investigating new ways of processing advanced materials. "This machine fits a niche between pure research and more industrially oriented research," says Hywel Price, the current director of the NSF who will take over as director of the new research unit.

The future of the tower housing the NSF is still under discussion. The highly visible building met with such local opposition at first that the SERC agreed to pull it down at the end of the NSF's working life; it is now something of a local landmark, and the destruction order has been lifted. Mercury, a telecommunications company, plans to mount equipment on it for receiving and transmitting radio signals, and the SERC is thinking about turning it into a science visitor centre that, among other things, would try to interest young people in careers in science and technology.

The irony has not been lost on the scientists working on the NSF, who are trying to do as much research as possible before the ion beam is shut off next week. "The closure of the NSF marks the end of an era for both Daresbury and for nuclear structure research in the UK", says laboratory director Alan Leadbetter.

**David Dickson**

## Life after death for nuclear physics

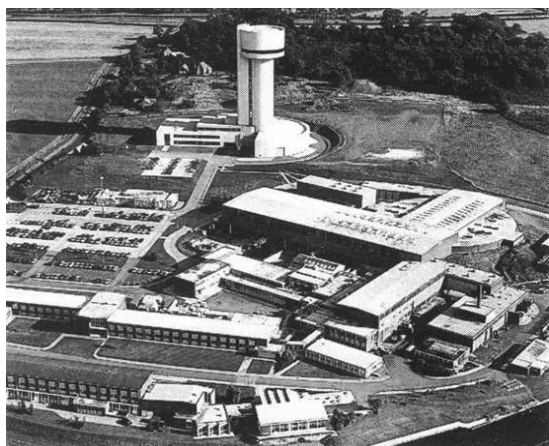
**London.** Nuclear structure physics in Britain may be down (see left), but it is far from out. Last month, the Science and Engineering Research Council approved £1.7 million (US\$2.5 million) for a new ion-source test-bed facility to be attached to the spallation neutron source ISIS at the council's Rutherford Appleton Laboratory. British physicists view it as the first step in a campaign to secure a radioactive nuclear beam facility, Europe's next major tool for research into nuclear structure.

At present, the amount of information obtained by machines designed to study the structure of atomic nuclei — such as the National Structure Facility (NSF) at Daresbury — is limited by the need to confine attention to naturally occurring nuclei. Studies are now planned at the European Laboratory for Particle Physics (CERN) of ways to add a 'post-accelerator' to the isotope separator ISOLDE in order to produce a beam of unstable radioactive ions, although still at relatively low energies. The experiments at the Rutherford Appleton Laboratory, to which a group of scientists from the NSF will now be attached, will complement these efforts by looking at techniques for producing proton beams at much higher energies and targets to withstand the intensities involved.

Using a spallation source to produce the radioactive ions may not turn out to be the best way forward; another approach is being explored at the heavy-ion accelerator GANIL in France. But success would give Britain a leg up in the competition for a machine, costing roughly £50 million, that has recently been given top priority by the European coordinating committee for nuclear physics, NuEPP. (The US Department of Energy has already asked for proposals from its laboratories for a similar machine.)

"The closure [of the NSF] was an act of folly and vandalism, but it still leaves our long-term future intact", says William Gelletly, formerly head of the NSF and now professor of physics at the University of Surrey. Gelletly is urging the SERC to persuade its European partners to locate the new machine in Britain as a way of regaining a central place in nuclear structure research that is being lost by the closure of the NSF.

**D.D.**



**Landmark tower outlives its usefulness.**

years ago at its funding agency, the Science and Engineering Research Council (SERC).

For about 25 members of the NSF staff, next week's final run will mean early retirement from a machine opened in the mid-1970s that had been expected to operate until late in the decade. For a fortunate few, a new ion-beam test facility at the Rutherford Appleton Laboratory (see right) or other projects at Daresbury will offer continuing employment.

Last year, proposals were submitted to the SERC for three projects to use the skills and equipment made available by the closure of the NSF. One was a facility to study small clusters of atoms or molecules, either in gases or attached to surfaces. A second was to build a much smaller Van de Graaff machine, capable of reaching energies up to 8 KeV, with potential applications as a mass spectrometer in fields ranging from archaeology to medicine. The third was to use the ion-beam injector to carry out crystallography studies.

Only the last has been successful. SERC committees decided that a central facility was not needed for the cluster project and that such work should remain in university departments. The proposal for an accelerator mass spectrometer was an interdisciplinary effort that fell between the responsibilities of the different research councils. Some scientific respectability has been salvaged by the SERC's decision to support



# MITI reshuffles institutes to keep up with the times

**Tokyo.** The Agency of Industrial Science and Technology (AIST), the main research component of Japan's Ministry of International Trade and Industry (MITI), last month implemented the first major reorganization in decades of four of its nine research institutes in Tsukuba science city, northeast of Tokyo.

The four institutes — the National Chemical Laboratory for Industry, the Fermentation Research Institute, the Research Institute for Polymers and Science and the Industrial Products Research Institute — have been combined into three: one for materials science, another for bioscience and the third a new style of interdisciplinary research institute intended to increase the flexibility of the ministry's research system. The interdisciplinary research institute will play an important role in a new ¥25 billion (US\$215 million) government-industry project to develop nanotechnology (see below).

The four institutes were created before the Second World War and were moved to Tsukuba in the 1970s. Their original research goals, as described by their names, are no longer relevant to trends in science and technology, leading some researchers in Tsukuba to refer to them as 'dinosaurs'.

In addition to restructuring completely the institutes' research departments, MITI has re-written Japan's laws to make it easier for the institutes to employ outside researchers from universities, companies and overseas on temporary contracts or for co-operative research. More than a third of the 1,100 or so researchers at the new institutes come from outside the institutes (see table).

The National Institute for Advanced Interdisciplinary Research is expected to

employ a flexible management system rather than the traditional rigid structure of national laboratories. The institute will carry out research on the basis of individual projects, and the projects will be reviewed every 3–5 years by outsiders, including foreign scientists.

The extended review process is controversial, but the agency is determined to introduce it and hopes to extend it to other AIST institutes. If successful, the institute will be the first national institute to follow in the footsteps of Tokyo University, which recently carried out an external review of its physics department by a panel of distinguished foreign and Japanese scientists (see *Nature* 360, 403; 1992).

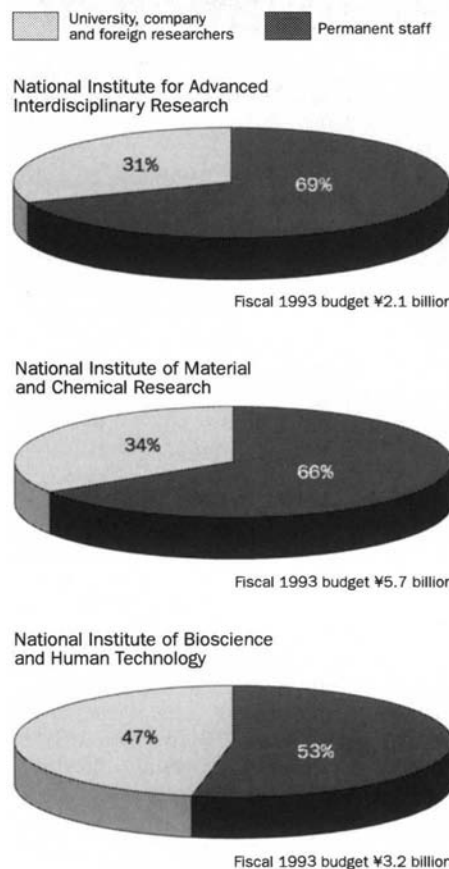
In its first year of operation, the interdisciplinary institute will concentrate on nanotechnology, cluster science and bionic design. That focus is expected to be broadened in the future.

The researchers in the three new institutes are now "spread like a mosaic" through the four buildings of the four former institutes, says Osamu Moriya, deputy director of AIST's general coordination division. Starting next year, members of the National Institute of Bioscience and Human Technology will be housed in the former Fermentation Research Institute as well as in a nearby building under construction.

Researchers at the National Institute of Material and Chemical Research will occupy the former National Chemical Laboratory and the adjacent building of the former Research Institute for Polymers and Textiles, which will have a new wing connected by a passageway to the chemical laboratory building. And scientists at the interdisciplinary

institute will use the former Industrial Products Research Institute. More than a hundred researchers will be affected by the moves, and construction will cost several billion yen.

Moriya says that the interdisciplinary institute will eventually double its size, to 70–100 permanent staff. Nearly ¥1,400 million (US\$12 million) has already been spent on new



equipment with money from the supplementary budget introduced by the government last October to fight Japan's recession. Starting this autumn, the institute will also be host to researchers from companies participating in the atom technology project.

**David Swinbanks**

## French voters shun 'greens'

**Paris.** The major surprise of last Sunday's first round of the French parliamentary elections was the unexpectedly poor showing of the greens, who won only 7.8 per cent of the vote. Last year, the Génération Ecologie/Verts alliance took 14.7 per cent in regional elections and were expected to win dozens of seats in the 577-seat National Assembly. Although their tally is larger than the 0.4 per cent that they received in the last parliamentary election five years ago, only three ecologist candidates have gone through to the second and decisive round of voting this Sunday, and only one may win a seat.

The fault appears to lie with an ambiguous response to an offer by leading Socialist Michel Rocard to help form a new wider centre-left party, which diminished voter confidence, as well as to a lacklustre campaign and a low turnout, barely two-thirds of the electorate.

**Declan Butler**

## It takes two to nano

**Tokyo.** US and Japanese competitors in the electronics and computer industries have joined forces for the first time in a research association of 30 major companies established last month to support MITI's nanotechnology project at the new National Institute for Advanced Interdisciplinary Research in Tsukuba (see above).

The US participants are semiconductor manufacturers Texas Instruments and Motorola, the Japanese subsidiary of the chemical company DuPont, and Yokogawa Hewlett Packard, a joint company of the US computer maker and Yakogawa Electric. Together with Japanese competitors such as Toshiba, Fujitsu, Hitachi, Sony and NEC, the companies have formed a research association called the Angstrom Technology Partnership (see *Nature* 360, 500; 1992).

The companies this year will contribute ¥200 million (US\$1.6 million) to the partnership, which is getting ¥500 million from MITI. Starting in October, each company will send one or two researchers to the Tsukuba institute.

**D.S.**

# Renovation of Paris museum leaves collection in disarray

**Paris.** The reopening later this year of the Grande Galerie of the Natural History Museum in Paris after almost 30 years should greatly contribute to the public's understanding of evolutionary biology. But a better showcase for some of the Galerie's 76 million specimens is a mixed blessing to researchers angry about the government's unwillingness to repair extensive damage to the museum's scientific collections from repeated flooding of an underground storage facility.

An architectural contemporary of the Eiffel Tower, the Grande Galerie was opened in 1889 to store and exhibit the museum's burgeoning collections. It was closed to the public in 1965 after falling into disrepair. (Its 600,000 litres of alcohol in specimen jars were also a fire risk.) In 1988, President François Mitterrand made an impromptu visit and, shocked by its state of ruin, came to the rescue by taking FF400 million (US\$80 million) from the government's prestige-projects fund. Five years later, the restoration is nearly done.

The large sum being spent on the Galerie contrasts with the emergency FF1 million allocated for restoring the collections, which have deteriorated more during the past seven years than in the previous three centuries. In 1986, the collections were transferred from the Galerie into an enormous underground "Zoothèque", built at a cost of FF45 million. But poor design (at a depth of 13 metres the Zoothèque is below the water table) and faulty construction have exposed the collection to a series of floods.

In addition to the need for immediate

preservation, there is also concern over the long-term fate of the collections. The museum is repairing the Zoothèque and suing the contractors, but staff would prefer a new facility to be built above ground. It is thought that an underground site was chosen to avoid the prohibition against increasing the amount of laboratory space in central Paris, part of an effort to decentralize research (see *Nature* 356: 373, 1992).

André Doumenc, a biology professor who studies marine invertebrates and molluscs, says that his recent request to attach an above-ground storage facility to the Grande Galerie was rejected because of political pressure to finish construction in time for the museum's bicentenary later this year. Researchers are concerned that the collections will be scattered throughout the provinces if no permanent home exists for them.

The new Galerie is in itself controversial, with many of the museum's senior academics criticizing it as a EuroDisney of evolution. But Michel van Praët, a marine invertebrate zoologist who heads the committee of scientists, architects and media experts responsible for the design of the new Galerie, says that it would be disastrous to give scientists complete control of public exhibitions because "they neither understand nor like the roles of the architects and media".

Van Praët has tried to make the museum more accessible; the results are imaginative exhibitions of the sweep of biodiversity, the mechanism of evolution and the effect of humanity on his planet, not to mention laser shows and sound effects.

Moreover, some of the criticism of the new Galerie can be traced to unhappiness over administrative reforms that have reduced the power of senior scientists. Last November, for example, a century-old system by which the museum's 26 professors controlled the museum's budget of FF140 million and 1,400 staff was abolished. In its place is a system more like that of a university, including a larger governing council and the replacement of life chairs by four-year appointments, renewable once.

Doumenc concedes that reform was overdue, but he is concerned that short-term appointments and separation of the museum's exhibition and research functions will dilute the commitment to doing science at one site.

**Declan Butler**

IMAGE  
UNAVAILABLE  
FOR COPYRIGHT  
REASONS

This 3.5-million-year-old *Australopithecus afarensis* couple is part of the new Hall of Human Biology and Evolution opening on 23 April at the American Museum of Natural History in New York. A successor to the popular Hall of the Biology of Man, which closed in 1984 after 25 years, the new permanent exhibition explores evolution, biology and anatomy as well as the origins of creativity, architecture and culture. The figures are part of a life-size diorama created on the basis of bipedal footprints discovered in Tanzania.

American Museum of Natural History

## Syria prodded to end human rights abuses

**Washington.** The US National Academy of Sciences has issued a report on human rights violations in Syria as part of a campaign to convince the Syrian government that restoring human rights is a precondition to scientific cooperation.

Twelve years in the making, a report by the academy's Committee on Human Rights documents the cases of 287 scientists, engineers and health professionals who have been detained or imprisoned without charge or trial in the past decade. Forty-nine appear to have been released since 1991 as part of an amnesty which may have freed 3,500 people, almost half of those believed to be held illegally. But the report warns that many of those released have been unable to resume their careers and that new or repeat arrests are common.

The academy last month presented the report to the Syrian ambassador in Washington, Walid Al-Moualem, and asked Syria to abide by the international human rights treaties it has signed, to release a list of detained and incarcerated scientists and to restore the autonomy of scientific and professional organizations.

An academy delegation was invited to Syria for a science meeting in November 1992, but Syrian President Hafez al-Assad refused to allow its members to meet government officials to discuss human rights issues. In response, the academy withdrew its participation. Nevertheless, the committee believes that the collapse of the Soviet Union has deprived Syria of its chief ally and benefactor.

**Jenna Roberts**

IMAGE  
UNAVAILABLE  
FOR COPYRIGHT  
REASONS

The renovation of the Grand Galerie, shown here as work began in 1991, is almost complete.



# Germany avoids large projects in favour of economic payoffs

**Bonn.** Germany's new research minister, Matthias Wissmann, has announced a strategy that closes the door on new national space missions, large-scale research projects and high-definition television but increases support for research in eastern Germany. Although his announcement last week contained no surprises, it lays to rest lingering uncertainties inherited from his predecessor, Heinz Riesenhuber.

Wissmann, on the job for only six weeks, has a short time to prove himself before the next general election at the end of 1994. Europe's biggest economy is in trouble, and German research, once the most generously funded in Europe, is being asked to produce more economically relevant results for less money. Riesenhuber had set the stage for streamlining research, but his decade of service made it hard for him to make the necessary changes. Wissmann, a lawyer, has no such loyalties. Space research is a luxury Germany can no longer afford, and high definition television has proved to be a white elephant. Wissmann raised no eyebrows by cutting these programmes.

The new strategy includes a greater commitment to environmental research, including alternative sources of energy now that nuclear energy is no longer politically acceptable, and a strengthening of modern technologies. Wissmann is singling out biotechnology in an attempt to reverse an unfavourable climate for genetic engineering.

National research centres created to develop nuclear energy, such as those at Karlsruhe and Jülich, will come under greater pressure to redirect their efforts. Wissmann is no fan of the large-scale centres, which are held to be inefficient and inflexible; he wants to see more competition within them and closer collaboration with industry. At the same time, he wants industry to increase its share of the national research budget, which has dropped from 62.3 per cent in 1989 to 58.4 per cent in 1992.

Wissmann's aims will not be easy to realize in the short term, but one goal is to strengthen dialogue between industry and academic institutions. In an attempt to break through research inertia (what he refers to as "decrepit structures") he plans to set up "innovation colleges" in eastern Germany, where researchers from both sides of the divide work for a few years on a defined government funded project, then take back their experiences to each domain.

Wissmann also wants to establish regular meetings of an advisory group made up of six scientists and six industrialists and supervised by the research ministry. The group is not intended to "tell industry what to do" in the manner of Japan's Ministry of International Trade and Industry but instead to develop a German model of cooperation and consensus.

**Alison Abbott**

# Swiss do well on patents but demand results

**Basel.** Switzerland performs better in biomedical sciences and physics than most countries and its institutions lead the world in the amount of research needed to obtain a patent, according to two recently published reports on Swiss science. But success is not uniform, and limited resources may mean that funding decisions will actually penalize successful researchers.

The reports are to be used this summer by the federal government to decide science policy for the next budget period 1996–99. The report from the science advisory body, the Swiss Wissenschaftsrat, shows that Swiss university researchers perform excellently according to bibliographic analysis. Citation rates in physics and biomedical research are comparatively high, with 64 per cent and 61 per cent of papers in those fields being cited by other scientists at least once. At the same time, the rate for clinical medicine is only 43 per cent and for mathematics it is 34 per cent.

In addition, the leading role played by industry in supporting research has not stifled the dissemination of results. "While

industry pays for 75–80 per cent of Swiss scientific research, 70 per cent of the publications come from universities", says François Da Pozzo of the advisory council.

The figures in the report will be used to grade performance and to determine funding, with success a possible indication that generous funding is no longer necessary. The result may be less money for capital-intensive fields such as physics and more for areas that, although weak, are deemed to be of national importance.

A new study from the government's economic agency, the Bundesamt für Konjunkturfragen (BfK), says that there should also be more cooperation between researchers and industry. Switzerland has more patents per inhabitant than any other country and acquires them very efficiently. A patent arises on average from research spending of only US\$1.25 million, exactly half the average for the European Communities. But the report warns that the future may not be so bright because Swiss high technology lags behind its international competitors in several fields.

**Oliver Klaffke**

# Centocor takes the bad news with the good

**Washington.** Last week was bitter-sweet for the biotechnology company Centocor Inc. of Malvern, Pennsylvania. Concurrent with an announcement that the company was stopping clinical trials of its leading drug, Centoxin (or HA-1A), preliminary data from a phase-III clinical trial of a second drug, CentoRx, indicated that the drug may help to reduce complications from blood clots associated with high-risk coronary angioplasty.

The company suspended the Centoxin trials in January after an interim analysis of trial data showed that the mortality rate among patients treated with the drug who did not have Gram-negative bacteraemia exceeded the rate among patients in the same group who received a placebo (see *Nature* 361, 290; 1992). The news caused share prices to fall by just 25 cents, to \$7.25. Most stock analysts had already written off Centoxin's chances of being approved by the US Food and Drug Administration (FDA) after the agency asked Centocor last April to undertake a second phase-III clinical trial because results from the first trial had failed to prove that the drug was effective. Centoxin is a human monoclonal antibody intended for the treatment of sepsis caused by Gram-negative bacteria.

An initial analysis of results from a phase-III clinical trial of the cardiovascular agent CentoRx indicated that the incidence of blood-clot-related complications was 35 per cent lower among patients treated with the drug than among those who received a placebo. Although this news was a relief for company officials, who say they plan later this year to file a product licence application for CentoRx with the FDA, it is Eli Lilly & Company of Indianapolis, Indiana, and not Centocor that is likely to reap most of the financial rewards.

Lilly entered into a strategic alliance with Centocor last July and, in return for an investment of \$100 million, obtained the marketing rights to Centoxin and an option to obtain the rights to CentoRx for an additional \$25 million if Centoxin was approved by 1 January 1994. With that outcome highly unlikely (the company has said that it will proceed with Centoxin if a way can be found to preselect patients who might benefit from the drug), Lilly stands to acquire the rights to CentoRx for nothing.

**Diane Gershon**

## Correction

The annual turnover of US Biochemical was incorrectly reported in a recent story about its pending sale to Amersham International plc (*Nature* 361, 199; 1993). The correct figure for last year is US\$38 million.



# Waldegrave says White Paper will not seek radical changes

**London.** The forthcoming White Paper (policy document) on British science and technology is likely to be "evolutionary rather than revolutionary", William Waldegrave, the cabinet minister responsible for science, last week told a meeting at Imperial College, London, organized by Save British Science.

The White Paper is still more than two months away, and Waldegrave refused to describe it in detail. But he lifted the edge of the curtain sufficiently to suggest that he is unlikely to adopt some of the more radical proposals submitted to the Office of Science and Technology last autumn, for example those from the Advisory Council on Science and Technology (ACOST) for a clear split between 'curiosity-driven' and 'mission-orientated' research councils.

While the ACOST report called for a sharper demarcation between the roles of government, the research councils and the scientific community, Waldegrave wants to ensure that all were pulling in the same direction. "We need somewhat more of a coherent national view about the broad direction of science and technology as they are relevant to this country, and a greater and more formalized exchange between players across the board", he said.

The principal actors in medical research, namely the research councils, charities and the health service, interact well, he said, adding that he "would like to see that more formally replicated in other areas". His words will reassure the Medical Research Council, which is concerned that implementation of the ACOST proposals could force it to hive off its research institutes through a process of "privatization".

Given the latest projections on public spending, it was no surprise that Waldegrave had little to offer those requesting a massive injection of new funds. Claiming that his hands are tied by the Treasury, Waldegrave said the real question was whether Britain was spending money in the right places.

A new report from Save British Science shows that Britain is continuing to fall behind its main economic competitors in terms of spending in research and development. But Waldegrave said he had "a great deal of scepticism about simplistic comparisons [of gross national products]". It was clear, he said, that economic performance is not automatically increased by spending more on science and technology. "If that were true, the states which make up the former Soviet Union would be in a strong position today", he said.

A consensus on the areas of science and technology on which the country should focus its resources would also provide a

framework for the educational system. "We have to send signals to people in science at schools and institutions of higher education if they are going to respond in a way which is good over a period of time."

There was a broad hint in the minister's speech that the White Paper is likely to contain measures to encourage the private sector to invest more in research. Waldegrave admitted that it was not doing enough at present but that it has been successful in areas — for example the pharmaceutical industry — where there has been a heavy investment. "We have to consider whether there are governmental things we can do to help this", he said.

He also held out some words of comfort for universities who feared that the White Paper will encourage them to turn away from fundamental research. There is no simplistic view that applied research is good and basic research is bad, he said. What should matter for universities is first that the work is excellent and interesting and second that it supports their educational role.

David Dickson

## Cold Spring Harbour promotes Stillman

**Cold Spring Harbor, New York.** Bruce Stillman has been named director of Cold Spring Harbor Laboratory, succeeding James Watson, who will become president on 1 January 1994.

"There are many things to be done here, and one person cannot do them all", says Stillman about the dual positions.

Stillman, 40, joined the laboratory in 1979 as a postdoctoral fellow and has been assistant director

IMAGE  
UNAVAILABLE  
FOR  
COPYRIGHT  
REASONS

**Bruce Stillman**

for the past two years. An expert on DNA replication, he will remain an active scientist as well as managing the laboratory's science programme. Watson, who has been director for 25 years and who turns 65 this year, will coordinate policy in a position created this month by the board of trustees. He will also be on sabbatical next year as a visiting professor at the University of Oxford.

The laboratory has 220 scientists and an annual operating budget of \$33 million.

Jeffrey Mervis

## CHINA IN BRIEF

**Beijing.** China is coping with a scientific 'talent fault' caused by the Cultural Revolution by promoting record numbers of scientists under the age of 50 at the nation's leading universities and by making research grants from an "outstanding young talent fund". Beijing University this year promoted 54 lecturers — 31 to full professor and 23 to associate professor — under the age of 50, a record that university officials say is "an exception" to the normal age limits for professorships. Some two dozen are under the age of 35, with the youngest, chemistry professor Lai Lu-Hua, only 29.

The rapid promotion of younger faculty can be disruptive in a country where seniority is still paramount. But it has become necessary given the fact that faculty must retire at the age of 65 and an increased emphasis on commercialization is drawing many younger professors into industry.

The National Natural Science Foundation received almost 200 applications for its first-ever round of young talent awards, which provide as much as 300,000 yuan (US\$55,000) over three years. Eleven of the 14 recipients have studied abroad, a fact that Chinese officials see as evidence of the value of international scientific exchanges.

■ Only 15 per cent of the more than 50,000 researchers at the various institutes operated by the Chinese Academy of Sciences will be allowed to continue with basic research in a sweeping reorganization of the scientific work force. Another 15 per cent will be given work collecting and analyzing data relating to natural resources and environmental management, while the rest — more than 30,000 researchers — will be left to find outside support for applied research that directly benefits the Chinese economy. The academy says that those remaining in basic research will receive preferential treatment and generous funding as part of an attempt to develop an elite corps of world-class scientists. Those going into applied research are expected to join such new creations as engineering centres, joint ventures and industrial conglomerates, aided by half a dozen shareholding companies being set up by the state.

■ China last week celebrated its first National Day for Science as part of the country's campaign to make scientific and technological research more market-orientated. A national fair in Beijing featured research institutions displaying their products to potential customers in an effort to increase the rate — now at 15 per cent — at which the country's current 20,000 research projects now result in commercial products. The date marks the fifteenth anniversary of a speech by senior leader Deng Xiao Ping that emphasizes the importance of science and technology to economic development.

Yau Qin Li

# Biologists denounce UC role in funding start-up companies

**San Francisco.** Some of the most eminent biologists in the University of California (UC) system, including two Nobel laureates, believe that a university plan to fund start-up companies to exploit faculty inventions would be an expensive mistake.

In a letter sent on 8 March to UC president Jack W. Peltason, 43 leaders in molecular biology said the plan could seriously hurt the teaching, research and public service functions of the nine-campus system. They said existing technology transfer programmes have been successful and that there is no need for a complex, expensive new structure. Last week, Peltason told the UC board of regents that he will address faculty concerns before moving ahead.

The letter gave voice to growing faculty concern over a proposal from Peltason to boost UC royalties by setting up a for-profit Technology Development Company to develop faculty ideas into prototypes (see *Nature* 360, 701; 1992). The company, of which the university would own 51 per cent, would then license those products or form companies to commercialize them.

Many universities have aggressive technology transfer offices and take equity shares

in start-ups, but few actually help to fund those ventures. The University of Chicago and Johns Hopkins University are among those that provide venture capital to prospective enterprises.

The University of California, operating in the heart of the \$6-billion US biotechnology industry, leads the nation's colleges in royalty income. It collected \$29 million for fiscal year 1992 and expects that figure to grow to \$40 million in the fiscal year ending this June.

Peltason sees the new corporation as a way to increase university revenues and to improve the state's sluggish economy, predicting that the expanded programme could produce \$222 million in gross royalty income by 2001 on operating costs of about \$2 million a year. The state, which normally claims about 15 per cent of the university's royalties, has promised to contribute its share to the enterprise.

But those who signed the letter say that the plan could end up costing the university money as well as diverting attention and resources from education. The plan would not change the inventors' share of royalty income but would centralize the technology

transfer function. A separate proposal would give less royalty money to inventors but more to their departments and their own departmental accounts.

Paul Boyer, professor emeritus and retired founding director of UCLA's Molecular Biology Institute, who drafted the letter, says that the plan could undermine the university's traditional goals. He is concerned about a potential decrease in the authority of individual campuses, insufficient contact with faculty inventors and a loss of government and private support for university research.

"We don't need a university bureaucracy to do what industry already knows how to do", said Harold Varmus, who signed the letter along with J. Michael Bishop, with whom he shared the 1989 Nobel prize for medicine. "It suggests that the university is much more interested in making money than it ought to be."

Those who signed the letter are not opposed to commercializing discoveries made in UC laboratories — Varmus is a paid adviser to two biotechnology companies in California, Gilead and Onyx, while two others, Edward Penhoet and William Rutter, are respectively chief executive and chairman of Chiron Corporation. But they see the new UC company as blurring the distinction between the two sectors. "This is not a modest proposal", says Penhoet. "This is a proposal to get seriously involved in commercial business." **Sally Lehrman**

## US meat inspections may get a strong dose of science

**Washington.** The Clinton administration has promised to introduce modern microbiology into the \$500 million-a-year US Food Safety and Inspection Service (FSIS) in response to recent food scares, but it is not clear that the proposed reforms will proceed quickly enough to ensure effective monitoring and control of harmful microorganisms in the food chain.

The US agriculture secretary, Mike Espy, and the FSIS administrator, Russell Cross, have committed themselves to a two-stage programme that will first upgrade existing inspection procedures while preparing a science-based approach to counter recent outbreaks of food-poisoning, the most recent leading to the deaths this winter of three children in the northwestern United States after eating hamburgers contaminated by *Escherichia coli* 0157:H7.

The adoption of modern scientific methods would require a major transfer of resources within FSIS from the 7,500 inspectors towards research into both the application of existing monitoring techniques and the development of new ones. Major elements of the science-based programme would include the identification of critical control points in the food chain and

the application of the latest advances in molecular biology, bioluminescence and biosensing. The programme, which was last week presented to a US House of Representatives agriculture subcommittee, would also set a microbiological baseline for

seven-year timetable for implementation has led to concern that the agency may drag its feet in reforming an existing system, based on visual inspection of meat, that all agree is obsolete.

Congress intends to expedite matters by convening a meeting of interested parties, including scientists, federal agencies and consumer groups, working under the supervision of a third party that would draw up binding proposals for reform. John McCutcheon, assistant administrator of FSIS, says that he "endorses the idea of such a public get-together" but denies that such proposals are a criticism of the FSIS's own reform plans.

Prominent food scientists have expressed optimism that the FSIS proposals will lead to substantive change. "The administration is committed to these changes and that is the crux of the issue", says Michael Doyle, professor of food microbiology at the University of Georgia and a member of the Food and Nutrition Board of the US National Academy of Sciences that has reported frequently on the issue. Doyle estimates that the plan could mean tens of millions of dollars a year in additional research funds. **Colin Macilwain**

IMAGE  
UNAVAILABLE  
FOR COPYRIGHT  
REASONS

**US inspectors check meat visually.**

different classes of animals and would monitor the performance of slaughterhouses and food plants against that baseline.

But the fate of the programme is uncertain. FSIS says that the cost of microbial testing for 20 per cent of the meat supply could be as high as \$58 billion a year, and its



# Wrong site for Japanese labs

SIR — David Swinbanks' article on the relocation of a Japanese laboratory (*Nature* 358, 703; 1992) mentioned my name but did not explain my stance and the background to the local and national protests. Allow me to explain the reasons why the location and experiments of Japan's National Institute of Health (NIH) in a densely populated residential area have been criticized by the mass media and public opinion in Japan.

The NIH is Japan's largest biotechnology and pathogen laboratory, with seven P3-laboratories and a large laboratory-animal facility with thousands of animals; it also uses organic solvents and large amounts of radioisotopes. The new NIH laboratory's seven-storey building (on a small 2,000-m<sup>2</sup> site) is immediately adjacent to houses, two welfare facilities for handicapped people, a major hospital and Waseda University with 40,000 students. So the NIH is a densely concentrated laboratory in one of the most densely populated areas in the world.

Since 1987, many residents have strongly opposed the transfer of the NIH to the new site. But in 1988, the NIH arrogantly started construction, backed by riot police. Although a lawsuit against the government has not yet been adjudicated, the NIH has recently moved to the new site and started experiments. We are, of course, still asking it to halt the experiments and to move to another, unpopulated site.

We contend that such a densely populated area is the wrong site for such a laboratory, because no biotechnology facility can be entirely safe. Furthermore, unknown genetically modified pathogens included in exhaust air from the laboratory are too dangerous, not only for the nearby residents but also for the population of Tokyo.

The rooftop vents of the NIH's seven P-3 laboratories, lab-animal facility and radioisotope laboratories are only 20 metres from houses and other neighbouring buildings. The NIH argues that it can set up such laboratories anywhere irrespective of environmental and population conditions and that the exhaust from the rooftop vents is clean and safe. If that is the case, why doesn't the NIH recycle the exhaust into the institute and let the staff themselves breathe it? The truth is, of course, that even the highly efficient filters cannot always catch all viruses in the exhaust from the P3-laboratory, and there is no safety standard yet for leaked recombinant DNA, tumour viruses and radioisotopes. The World Health Organisation's *Laboratory Biosafety Manual* (Geneva, 1983) states: "The exhaust air from the

laboratory should be discharged directly to the outside or the buildings exhaust system so that it is dispersed away from occupied buildings and air intakes" (p. 26, emphasis added). But just outside the NIH are tens of thousands of residents.

The hitherto accepted guidelines for biotechnological experiments have referred only to laboratory conditions, not to the neighbouring environment. However, concerned governments have begun to pay attention to environmental aspects of laboratories. The *Directive on the Contained Use of Genetically Modified Micro-organisms* (23 April 1990) from the Council of the European Communities, reads:

The precise nature and scale of risks associated with genetically modified micro-organisms are not yet fully known and the risk involved must be assessed case by case;

It may be considered appropriate to consult the public on the contained use of genetically modified micro-organisms;

The user shall carry out a prior assessment of the contained uses as regards the risks to human health and the environment that they may incur." (*Official Journal of the European Communities*, No L 177, 8 May 1990.)

The NIH should apply the regulations proposed by the WHO and the EC to biotechnology laboratories in Japan and a comprehensive environmental impact assessment should be carried out.

If local residents were to be infected with genetically modified pathogens from the NIH, infection might spread not only to Tokyo but all over the world. Thus we are campaigning not only for Japanese but also for humankind as a whole.

## Shingo Shibata

(Chairman, Plaintiffs of Lawsuit against Japan's NIH),  
1-18-6 Toyama, Shinjuku,  
Tokyo 162, Japan

## Gene therapy

SIR — David Dickson reported recently (*Nature* 361, 387; 1992) on the approval by British authorities of the country's first gene therapy experiment, involving the transplantation of a functioning adenosine deaminase (ADA) gene into the bone marrow cells of a child suffering from ADA deficiency. In citing previous experiments carried out in the United States and Italy, Dickson writes that "in those cases, only T-cells were altered and the process had to be repeated at regular intervals". As the British protocol will involve manipulation of stem

cells as well as T cells, these experiments may instead "result in one of the first permanent implantations of a manipulated human gene".

The report is incorrect. The ADA-deficiency gene therapy clinical trial currently under way in Italy is based on retroviral vector-mediated transduction of both peripheral blood lymphocytes and bone marrow cells, according to a protocol approved by the Italian National Ethical Committee and by the Ethical Committee of the San Raffaele Research Institute in Milano, where the trial is being carried out in collaboration with the Pediatric Clinic of the University of Brescia. *Nature* correctly reported on this trial shortly after its beginning ("Italians first to use stem cells", by Alison Abbott, 356, 465; 1992), by mentioning that it is based on the use of two vectors, differing in one base pair, for the two tissues. This would, it is hoped, allow monitoring of the circulating ADA-positive T-lymphocytes and determine whether they derive from lymphocyte infection or from progenitor cells infected in the bone marrow. Furthermore, a Dutch research group headed by Dinko Valerio received approval in February 1992 for gene therapy of ADA deficiency by infection of bone marrow cells, although by a quite different protocol and this too was reported in *Nature* (355, 190; 1992).

Fulvio Mavillo

Claudio Bordignon

*Istituto Scientifico San Raffaele,  
Via Olgettina, 60, 20132 Milano, Italy*

## Hard physics

SIR — Readers of your article "Can quantum theory be understood?" (*Nature* 361, 493; 1993) may be interested to learn of a new research project set up to investigate A-level students' conceptions of quantum physics. Despite the fact that it is 70 years since concepts such as 'wave-particle duality' were first formulated, no significant research on students' understanding of such conceptually unusual topics has been carried out.

Many A-level physics students do experience considerable difficulties in understanding the quantum physics section of the syllabus. The study will address the following questions. How can student understanding be identified? What patterns are found in student understanding? What implications have these patterns for learning? How can effective teaching be achieved?

I would be most interested in your readers' reactions.

Azam Mashhadi

*Department of Educational Studies,  
University of Oxford,  
15 Norham Gardens, Oxford OX2 6PY, UK*



# No easy answers for UK science

David Dickson

**The British government is promising a national strategy for science and technology in its forthcoming White Paper on research. Will its bite live up to its bark? Or will it end up merely rearranging the deck-chairs on the Titanic?**

**London.** What would be the impact on academic scientists of a national strategy for British research? Will the government's chief scientist have sufficient clout to ensure that government departments work together — rather than against each other — in pursuing such a strategy? How can both the private sector and the Treasury be persuaded to provide the research money needed to enable the goals of the strategy to be met? And is tinkering with administrative reforms little more than rearranging the deck-chairs on the Titanic?

There were more questions than answers at a meeting organized by *Nature* at the Royal Society in London on 19 March to discuss the government's forthcoming White Paper (policy document) on research and innovation. Many questions are likely to remain, even after the White Paper is published; all the more reason, some suggested, for the government to consider publishing a document with "green edges" (a reference to the fact that government policy statements on important topics are usually preceded by a "Green Paper" in which preliminary ideas are thrown out for public debate).

But with government discussion on the White Paper now entering what Mr William Waldegrave, the cabinet minister responsible for science, describes as the "last furlong" — one publication date being quoted is 17 May, although Whitehall sources indicate that this appears optimistic — the broad outlines of its likely content are already becoming apparent (see page 282). Attention is now turning to the implications of their implementation.

The good news is that Britain seems likely to have an explicit policy for science, embedded within the broader framework of an industrial policy built around consensus on the main scientific and technological needs of the future. The market on its own, dominated as it is by financial institutions with a greater interest in short-term dividends than long-term investments, has failed to do this, as other countries have realized.

Indeed, substitute the concept 'critical technologies' for that of 'generic technologies' — the name given to new technologies which have applications in a range of industrial fields, support for which is rapidly becoming part of the accepted wisdom of Whitehall — and government thinking (bar the lack of extra funding) begins to look increasingly like that of the Clinton administration in the United States.

But, as last week's meeting revealed, two large areas of concern still remain. First, that

a single-minded concentration on 'wealth creation' as the principal driving force behind the government's thinking could eclipse other goals — such as the need to protect curiosity-driven science, to pay researchers a fair salary, or to take steps to help women pursue a career in science — that remain desirable in their own right, even if in the short term they do not contribute to economic growth in a cost-effective way.

Second, excessive concentration on administrative details must not obscure the



**Sir Mark Richmond**

larger picture, for example the fact that only 15 per cent of spending on research and development is directly controlled by the government, or that a major problem in the future will be a lack of high-quality science students. The danger here is that the bigger questions, and in particular the need to create a social and economic environment that encourages private investment in research by both individuals and institutions, may appear so intractable that rearranging the deck-chairs becomes a seductive diversion.

Sir Mark Richmond, chairman of the Science and Engineering Research Council (although emphasizing that he was speaking in a personal capacity) opened the meeting by pointing out that the need for national strategies for science and technology now seems widely accepted around the world. It is, for example, a stated objective of the new US administration; and Britain's colleagues (or competitors) in the European Communities are all moving to develop much more purposeful strategies for wealth creation based on a strong fundamental science base.

Until recently, said Richmond, the only policy that the British government appeared to have on the issues was to have no policy. "Now wealth creation, technology transfer and innovation are all seen as issues of great concern, and there is a widespread desire to

see the science base used for a purpose; or for it to have a 'mission', to use the developing jargon."

The central problem is how to develop and implement a national strategy without undermining the excellence and quality of a research base that depends on the imagination of scientists free of the constraints of bureaucracy. It is here that the research councils have a crucial role, and Richmond said that he hopes to see these bodies, functioning much as they do at present, reconfirmed in the White Paper as "the main vehicles for using public money to support research for wealth creation".

Michael Brady, professor of engineering at the University of Oxford and a member of the Advisory Council on Science and Technology (ACOST), said that there are a number of reasons why the council feels that the current system of research support should be changed. One need is for a clear separation between those who fund research and those who "provide" it. "We need to know what is proposed, why, and how the outputs are to be judged", he said.

Brady vigorously defended the controversial proposals made in the ACOST report to introduce different funding methods for curiosity-driven research and mission-orientated research. The first of these, he said, should be judged according to the criteria of excellence, timeliness and promise, with relatively little concern about whether such research would contribute directly to wealth creation.

It is this type of research, he said, that would be financed through the Council for the Advancement of Scientific Knowledge (CASK) which ACOST had proposed in its submission to the government on the White Paper. In contrast, mission-orientated research — funded by a separate research council operating through various programme directorates — should be driven by its relevance to clearly specified policies, and judged according to its relevance to pre-defined goals.

He accused ACOST's critics of misrepresenting the council's recommendations by assuming that the concept of curiosity-driven research covered all basic research. "Curiosity-driven research may turn out to be useful, and it may even be possible to assign a probability to this happening; but utility should not be the criterion by which such research is judged," he said. Similarly, it might be necessary to carry out some basic research as part of a mission-orientated project; "but I would not describe it as being

curiosity-driven", he said.

Brady denied that ACOST's proposal would threaten the future of world-class laboratories such as the Medical Research Council's (MRC's) Laboratory for Molecular Biology (LMB) at Cambridge. Such institutions would expect to build "long-term partnerships" with their main funding agencies, he said, which would keep their research programmes intact.

But, he added, there are a large number of government-funded laboratories that need to become much more relevant to national

needs. As a result, the privatization of the "vast majority" of government-funded research institutes is a "reasonable goal to aim for", he said.

Brady also denied that full implementation of the ACOST proposals would, as many critics have



**Dai Rees**

suggested, lead to a fragmentation of Britain's research effort. "We have to have the various parts interacting with each other", he said. "That is what we have tried to take steps towards creating with the proposals which we have made."

But despite his soothing words — including those about the MRC's Cambridge laboratory — ACOST's proposals were still firmly rejected by Dai Rees, secretary of the MRC. Rees complained of "false prophets preaching false remedies", some of whom spoke as if curiosity-driven research should be uncoupled from any practical benefit that it could bring. "To my mind, that is a cop-out", said Rees.

Even in areas such as the neurosciences and biotechnology, it is important to avoid an institutional tunnel vision that can protect special research initiatives from the challenges of other disciplines. Clinging to such protection is a recipe for "intellectual impoverishment", he said. A better approach, he suggested, is that adopted by the MRC, which combines vertical integration — linking together basic, applied and clinical research in a particular discipline — with horizontal competition. Using this method, said Rees, it is possible to compare the value for money obtained in one area of research against another.

Rees also stressed the need for flexibility in funding procedures. It is this that allows the MRC to align itself with the pharmaceutical industry for some of its activities, and with the food industry for others. "We have to be flexible enough to adopt different postures to different partners", he said. "This is why monolithic schemes which are seen as accommodating every need of every industry are doomed to failure."

Where change is needed is in recognition

of a significant change in the traditional relationship between universities and research councils. Previously, university scientists have been virtually able to write their own grant cheques and follow research topics of their choice. Now, financial stringency means research must be much more carefully planned, and the planning procedure has to be coupled with mechanisms for assessing the output of research grants to see whether research groups have delivered on their promises. "This is a major challenge to the research councils and the universities over the next five years", he said.

Rees's criticisms of ACOST were backed by Sir Aaron Klug, director of the LMB itself, who claimed that the council's desire to distinguish between the "purchasers" and the "providers" of research could not be maintained. "The MRC invests in, and does not 'purchase' research", he said, adding that there is no sin in having a council running its own research institutes, as long as these are subject to close scrutiny and review.

Further disagreement with the ACOST proposals came from Sir Eric Ash, rector of Imperial College, London. He suggested that the distinction between curiosity-driven and mission-orientated research is not only dangerous, but out of date. For several decades after the Second World War, the main progress in fundamental science came in fields such as high-energy physics which did not lead directly to applications.

These sciences were characterized by an interest in the underlying simplicities of the Universe. In contrast, recent years have seen the rapid growth in basic sciences such as molecular biology, chaos theory and nonlinear-thermodynamics, which are characterized by complexity — and where the chances of producing useful results are almost inevitable. "This perception cannot and should not be ignored by those entrusted with creating science policy", he said. "The separation advocated by ACOST would be wholly counterproductive."

Ash endorsed the need for a top-level science-policy advisory committee. But while those on the committee need to have "significant jobs" in industry or academic institutions, such individuals could often not devote the time required to carry out the committee's tasks adequately. The solution could be for such an advisory body to recruit two or three young active scientists who would be seconded to it for at least one day a week for three years.

Speaking from the floor of the meeting, Sir Hermann Bondi, a former chief scientist at the Department of Defence and head of the Natural Environment Research Council, defended basic science as the greatest spiritual adventure mankind has known. "We should be quite clear that there is an area of science which is supported without any view of wealth creation" he said.

At the same time, Bondi warned that in many fields of science, the cost of support-

ing a scientist adequately is rising faster than the funds likely to be made available. In such fields, faced with an apparent excess of creative talent, the number of scientists should be reduced, he said.

Fiona Steele, of the Confederation of British Industry, emphasized that it is important to address the "macro-issues" and not get bogged down in administrative details. "The main point is that we need a national strategy", she said. "The whole debate has to be set against a much wider canvas, of which science and technology form a part; rather than just advising the minister on a policy for science, we should address the question of how it can contribute to the country's broader goals."

A plea for more attention to the needs of science education came from Peter Hughes, head of the science department at Westminster School. Hughes said that he hoped that splitting science off from the responsibilities of the Department for Education "will not mean that science and education are not going to talk to one another". It is important to ensure that the department continues to get an adequate input of thinking about science, and that the Office of Science and Technology receives the right input about teaching.

Others stressed that a key factor leading to the weakness of research in British industry is the low level of pay to industrial researchers. "If industry is complaining about the lack of qualified manpower, it could start by paying them on the same levels as it

pays to its accountants," said Richard Jones, pro-vice chancellor of the University of East Anglia.

Finally, there was broad agreement that a key factor that needs addressing if the gap is to be closed between Britain's success in scientific



**Michael Brady**

discovery and its relative weakness in turning these discoveries into commercial products is the low level of technical and managerial expertise on the shop floor. "One of this country's major failings is the lack of an adequate supply of the technical and managerial skills needed to ensure that ideas are successfully turned into products", said Margaret Sharpe of the Science Policy Research Unit at the University of Sussex.

It was one more reminder to the government that there are no simple solutions — and that those that do exist (such as better technical and management training) are likely to require money. In this context, last week's budget, in which the government revealed that the pressure on public spending are unlikely to ease over the next three years, has provided little comfort. □



# Where the AIDS virus hides away

**Two research groups have independently shown that the long and variable latency period between HIV infection and overt AIDS is explained by replication of virus in the lymph nodes. Even Duesberg must pay attention.**

EVEN when walking in the dark, it helps to have a map. Then at least it may be possible to identify obstacles accidentally encountered. That is the spirit in which the two articles on the pathogenesis of AIDS in this week's issue (pp 355 & 359) should be welcomed. Neither can be counted as a step towards the more effective treatment of the disease; indeed, each tends to emphasize that the difficulties that lie ahead are formidable. But the two papers do explain why the time-lag between infection by HIV and the appearance of overt AIDS is both long and variable. In due course, this insight and the knowledge of how and where HIV is sequestered while its effects are latent cannot but be important.

The belief that AIDS is caused by the retrovirus called HIV (for human immunodeficiency virus) is widely but not generally accepted. Some, who usually refer to that proposition as the 'HIV hypothesis', hold that other agents are responsible. Professor Peter Duesberg, from the University of California at Berkeley, has been the cheer-leader of that group, first by drawing attention to some surprising anomalies of infection by HIV, such as the difficulty, early in infection, of recovering virus from the T lymphocytes whose decline usually accompanies the onset of AIDS, more recently by arguing (with little justification) that drug-taking is responsible for AIDS (see *Nature* 362, 103; 1993). His more cogent objections have now been incidentally answered.

That the natural history of a retrovirus should differ from that of the viruses most commonly infecting people is not, or should not have been, surprising. They are mostly DNA viruses, which ordinarily replicate within cells by hijacking the preexisting machinery of DNA transcription and translation. Then, the course of infection may be as predictable as with, for example, that by the measles or influenza viruses. The production of new virus particles depends on the activity of cells already infected; the proportion of susceptible cells infected then increases steadily as virus particles spread out from the foci of infection as, with a time-lag, does the proportion of lymphocytes carrying neutralizing antibodies.

The genomes of retroviruses, by contrast, consist of RNA. Those of the lentiviruses, of which HIV is one, come equipped with a gene specifying a reverse transcriptase (for converting RNA into the complementary DNA). Although the RNA genome may be used, as if it were one of the infected cell's own messenger molecules, to

generate the proteins that would allow an intact virus particle to be regenerated, by far the more efficient means of replication is that DNA complementary to the viral RNA should be incorporated in the genome of the cell, where it will indefinitely serve as a template for the production of its own genomic RNA and thus for intact viral particles.

Apart from AIDS, the only human disease unambiguously known to be caused by a retrovirus is adult T-cell leukaemia, common in southwest Japan, but it will be surprising if there are not others. It is also relevant to this thumbnail sketch of what textbooks now say about the mechanism of viral infection that some DNA viruses (hepatitis-B and Epstein-Barr, for example) also appear to be integrated in the genome of the cells they affect, even if perhaps only rarely.

Duesberg's puzzle about the pathogenesis of AIDS is that it is difficult to recover from 'helper' T lymphocytes, whose attrition for many patients signals the onset of overt AIDS, virus particles that might plausibly infect others. What kind of a virus is it that seems not to affect its chief targets? That has been the question. The answer, now made plain, is that the virus is alive and well in the lymph nodes, among other sites, of those infected with HIV.

The two groups which have independently demonstrated this have used the polymerase chain reaction (PCR) to look for signs of the presence of HIV in various types of cells of the lymphatic system of people infected with HIV. They have searched separately for the RNA and DNA forms of the viral genome. Bolognesi and Temin explain, on page 292, that such a search would not have been successful even ten years ago; the present applications of the technique make it possible to distinguish cells carrying HIV in its RNA or DNA forms from those that do not carry traces of the virus. Bolognesi and Temin also suggest how the strategy of AIDS research should be amended, in the light of the new findings, to yield a better understanding of the pathogenesis of AIDS.

With hindsight, it is not surprising that the lymph nodes (but also the spleen, adenoid glands and tonsils) should be sites at which the presence of HIV is most readily demonstrated; that, after all, is where most helper T cells usually reside. In the lymph nodes just as in the peripheral blood, cells carrying the DNA form of the HIV genome are more common than those marked by RNA; true latency seems to be the rule. The

big surprise is the degree to which virus seems to multiply in these organs. But it is also striking that there are heavy concentrations of viral particles in the spaces between the cells of some structures in the lymph nodes. Does their location there allow them to infect T cells in their passage through the glandular structures of the immune system?

Whatever the answer, the conclusion must be that the apparently latent period between infection with HIV and the overt symptoms of AIDS is not clinical latency at all, but rather a period in which the replication of the virus proceeds apace in the lymph nodes and some other tissues of the immune system. The suggestion by Bolognesi and Temin that there should now be more systematic investigations of the degree to which the RNA and DNA versions of the HIV genome occur in the various cells of the immune system at different stages in the development of the disease is plainly an important way of building on these discoveries. It may even suggest ways in which latency might be prolonged, in turn a better approximation to the notion of a cure for AIDS than anything else in sight.

These developments will be a conundrum for Duesberg and his disciples. Will they now abandon their opinion that the 'HIV hypothesis' is an insubstantial hypothesis, less plausible than others, that AIDS is caused by drug-taking for example? Duesberg's advocacy of his cause, and in particular the manner of his advocacy, has maddened the AIDS research community. So too has this journal, by occasionally venturing to suggest that the more cogent of Duesberg's objections deserved to be taken seriously, if only because his reputation in other aspects of virology deserves respect. But now he has a clear chance to demonstrate that he is indeed the open-minded fellow he has always claimed to be: will he repudiate his past position?

This matters not so much for Duesberg's standing among his colleagues, but for the interests of those with AIDS who follow him. For those infected with HIV, the 'HIV hypothesis' implies a terrible finality: you will die, but nobody can tell you when or even how. It is natural that other hypotheses about the pathogenesis of AIDS should be emotionally preferable. This week's developments suggest that, nevertheless, the alternatives are even less tenable than seemed likely a few years ago. Duesberg, having led many people with AIDS on a seductive path, should now admit the likelihood that he is mistaken. **John Maddox**

# Mammoths in miniature

A. M. Lister

THAT so quintessentially prehistoric a creature as the mammoth should have survived to coexist with the Egyptian pharaohs is a revelation reminiscent of the discovery of the first living coelacanth. That the remains in question should represent that ultimate palaeontological paradox — a population of dwarfed mammoths — further conspires to make the discoveries described by Sher and his colleagues on page 337 of this issue<sup>1</sup> one of the most extraordinary fossil finds of recent times.

The woolly mammoth (*Mammuthus primigenius*) died out on the continental land masses as part of the Late Pleistocene mass extinction of large mammals. Its latest radiocarbon-dated occurrences in Europe and northeast Siberia, North America, and north central Siberia — apparently the last stronghold — are about 12,000, 11,000 and 10,000 years ago respectively (ref. 2; L. Sulerzhitsky, personal communication). The new finds come from Wrangel Island in the Arctic Ocean, some 200 km off the coast of northeast Siberia, and a large series of radiocarbon dates, duplicated by two laboratories, indicates the presence of mammoths there between 7,000 and 4,000 years ago. That extends the species' survival by some 6,000 years beyond its previously supposed global extinction.

The Wrangel mammoths are bound to fuel the already vigorous debate about the causes of Late Pleistocene extinction. By 12,000 years ago, human populations had spread throughout the Eurasian continent, including the remote regions of northeast Siberia. At Dyuktai Cave, for example, dated between 14,000 and 12,000 years ago, mammoth bones occur together with spear points and knives<sup>3</sup>. Such finds do not in themselves constitute direct evidence of mammoth hunting. But proponents of the 'human overkill' model<sup>4</sup> are likely to regard them as consistent with hunting to extinction on the mainland, given that on Wrangel there is no evidence of human occupation at that time and the mammoths survived.

This, however, is not the explanation favoured by Sher and his colleagues. Like many others, they see large-mammal extinction as precipitated primarily by climatic and vegetational changes between 15,000 and 10,000 years ago. Warming of climate caused the previous rich, 'mosaic' vegetation of herbs and grasses (the steppe tundra) to be replaced by today's less diverse and productive zonation of boggy tundra and coniferous forest, and steppe-tundra

mammals, such as the mammoth, died out<sup>5</sup>. But whereas such changes transformed the mainland, the present-day vegetation of Wrangel is peculiar in that it seems to represent a relict of the original steppe-tundra biome<sup>6</sup>. There is no direct palaeobotanical evidence that the present environment of Wrangel has indeed been continuous back to the Late

moths to have survived on an isolated island, then it should come as no surprise that the animals were dwarfed. Size reduction of large mammals on islands was a common phenomenon in the Pleistocene, affecting populations of deer, hippo, elephantids and other taxa<sup>13</sup>. Examples of dwarfed-mammoth populations are few, comprising only *Mammuthus exilis* from the California Channel Islands<sup>14</sup>, and possibly *M. lamarmorae* from Sardinia<sup>15</sup>. The Wrangel animals were not the smallest of dwarf elephantids (that honour goes to *Elephas*

IMAGE  
UNAVAILABLE  
FOR COPYRIGHT  
REASONS

A romantic, but unrealistic, view of the woolly mammoth from 1909. Mammoths lived on a kind of prairie, now vanished, called steppe tundra, which was rarely covered in snow.

Pleistocene, but if it were, this could explain the late survival of mammoths there.

There is also no direct evidence on the final cause of extinction on Wrangel some 3,500–4,000 years ago. A climatic/vegetational cause is possible, but environmental conditions on the island at that time are unknown. The earliest known human occupants were palaeo-Eskimo hunting/fishing populations with radiocarbon dates of about 3,000 years ago<sup>7</sup>; if their forebears were already there a few hundred years earlier, they would have been contemporary with (and perhaps hunted) the mammoths. At Zhokhov Island, also in the Arctic Sea north of Siberia, hunters were present as early as 7,000–8,000 years ago, and used implements made of mammoth ivory<sup>8</sup>. If such people were responsible for the demise of the Wrangel mammoths, it would fit into the general pattern of historical decimation of relict island faunas at the hands of early settlers, for example in the extinction of lemurs on Madagascar, moas in New Zealand and pygmy hippos on Cyprus<sup>9–12</sup>.

If it is appropriate for the last mam-

*falconeri* of Malta, which was barely more than a metre high<sup>15</sup>), but they were perhaps only 180 cm high and weighed around 2 tonnes, compared to 320 cm and 6 tonnes for typical European *M. primigenius*.

One of the most satisfying aspects of the Wrangel discoveries is the occurrence on the island of normal-sized mammoth remains as well as the small ones, the large individuals giving radiocarbon dates of 12,000 years or older and representing the ancestral population. This correlation of size with age corroborates the validity of the very young dates for the dwarf remains, and indicates that dwarfing occurred after isolation of the island about 12,000 years ago. Particularly instructive is the comparison<sup>1</sup> of the Wrangel dwarfs with mammoth remains from Berelekh, on the adjacent mainland, which date to about 12,000 years ago. These animals were very small — not as small as those of Wrangel, but at the lowermost end of the 'normal' range of mainland woolly mammoth variation<sup>1,16</sup>. It is tempting to suppose that this reduction in size was associated with suboptimal habitat conditions immedi-



ately preceding mainland extinction.

To what extent can the island dwarfing phenomenon be seen as an extension of the same process? Island dwarfing has been subject to various interpretations, but generally they relate to resource limitation, lack of predation, or both<sup>17</sup>. Body-size reduction of mainland populations prior to extinction may be largely produced by direct environmental (ecophenotypic) effects on growth. This

mechanism may also be responsible for the beginnings of island dwarfing, though the latter often proceeds far beyond anything seen on the mainland, suggesting that other processes, including genetic change and developmental adjustment, are at work. □

A. M. Lister is in the Department of Biology, University College London, Gower Street, London WC1E 6BT, UK.

1. Vartanyan, S. L., Garutt, V. E. & Sher, A. V. *Nature* **362**, 337–340 (1993).
2. Stuart, A. J. *Biol. Rev.* **66**, 453–562 (1991).
3. Praslov, N. in *Late Quaternary Environments of the Soviet Union* (eds Velichko, A. A., Wright, H. E. Jr & Barnosky, C. W.) Ch. 29 (University of Minnesota Press, Minneapolis, 1984).
4. Martin, P. S. in *Quaternary Extinctions* (eds Martin, P. S. & Klein, R. G.) 354–403 (University of Arizona Press, Tucson, 1984).
5. Guthrie, R. D. in *Quaternary Extinctions* (eds Martin, P. S. & Klein, R. G.) 259–298 (University of Arizona Press, Tucson, 1984).
6. Yurtsev, B. A. in *Paleoecology of Beringia* (eds Hopkins, D. M., Matthews, J. V. Jr, Schweger, C. E. & Young, S. B.) 157–177 (Academic, New York, 1982).
7. Tein, T. S. In *New Archaeological Monuments in the Northern Far East* 53–63 (North-Eastern Complex Research Institute, Far Eastern Science Centre, USSR Academy of Sciences, Magadan, 1979).
8. Pitul'ko, V. & Makeyev, V. *Nature* **349**, 374 (1991).
9. Dewar, R. E. in *Quaternary Extinctions* (eds Martin, P. S. & Klein, R. G.) 574–593 (University of Arizona Press, Tucson, 1984).
10. Trotter, M. M. & McCulloch, B. in *Quaternary Extinctions* (eds Martin, P. S. & Klein, R. G.) 708–727 (University of Arizona Press, Tucson, 1984).
11. Diamond, J. M. *Nature* **359**, 15 (1992).
12. Simmons, A. H. *Nature* **333**, 554–557 (1988).
13. Sondaar, P. Y. in *Major Patterns in Vertebrate Evolution* (eds Hecht, M. K., Goody, P. C. & Hecht, B. M.) 671–707 (Plenum, New York, 1977).
14. Stock, C. & Furlong, E. L. *Science* **68**, 140–141 (1928).
15. Caloi, L., Kotsakis, T., Palombo, M. R. & Petronio, C. in *The Proboscidea: Trends in Evolution and Palaeoecology* (eds Shoshani, J. & Tassy, P.) (Oxford University Press, in the press).
16. Lister, A. M. & Joysey, K. A. in *Structure, Function and Evolution of Teeth* (eds Smith, P. & Tchernov, E.) 185–213 (Freund, Jerusalem, 1992).
17. Heaney, L. R. *Evolution* **32**, 29–44 (1978).

## OPTICAL COMPUTING

# Lighting up logic

Robert W. Keyes

THE thought that optics can be used to perform digital logic operations is not new — it has attracted interest since the invention of the laser. Although it has met with limited success, it has been continuously encouraged by the rapid advance of optical technology, exemplified by developments such as low-loss optical fibre, continuous-wave, room-temperature semiconductor lasers that operate at a variety of wavelengths, and quantum wells. Workers at the Optoelectronic Computing Systems Center of the University of Colorado now claim\* to have devised the first simple programmable digital optical computer.

The most successful attempts to use optics to do the real digital operations known from electronic computers have been found in systems that use electro-optic devices to do the work of making digits interact and influence one another, and use optical signals to move information from place to place. The first of these to be demonstrated was apparently that developed by B. K. Jenkins *et al.* of the University of Southern California (*Appl. Opt.* **23**, 3455–3464; 1984). The authors used a so-called liquid-crystal light valve, a combination of a photoconductor and a liquid crystal, as the electro-optic element that allowed a light input to influence a second light beam.

Light impinging on the photoconductor changed the electric field applied to the liquid crystal, changing the valve's opacity. The system performed a reasonably complicated sequence of operations, but the slow response of the liquid crystal severely limited its speed.

Much attention was attracted to the system announced in early 1990 by AT&T Bell Laboratories, which was based on a device called the SEED, for self-electro-optic-effect device (M. E. Prise *et al.* *Appl. Opt.* **30**, 2287–2296; 1991). The SEED uses photoconductivity induced in a semiconductor *pn* junction by an optical input to alter the electric field to which a thin layer in the junction is exposed. The change in electric field moves the position of an absorption line to which a laser has been tuned. Response times are more typical of semiconductor devices, orders of magnitude faster than liquid crystals. Both the University of Southern California and AT&T systems beamed signals through free space between devices.

The electro-optic device in the new computer is a lithium niobate switch, in

which an applied electric field can control the paths taken by beams of polarized light. The logic function performed is an interchange of beam paths in response to an input (the action of a Fredkin gate in conventional electronics is similar). Optical inputs to gates are converted to electrical signals by semiconductor diodes that can control the field applied to the lithium niobate. The machine is novel in its dependence on the transit times of signals to synchronize operation as a pipelined string of bits and in the use of optical fibres to route signals between devices. The fibre connections give it the flexibility in interconnection patterns needed to implement complex logic functions (R. W. Keyes *IEEE Circ. Dev. Mag.* **7**(3), 32–35; 1991). In contrast to the previous optical systems, which were programmed by physical structures that could not easily be changed, the new machine uses an optical delay line as a memory containing both instructions and data.

The need to use electro-optic devices to perform logic operations in these systems originates in basic physics, in an essential difference between light and electricity. Logic involves the interaction of two or more streams of data, and light interacts only weakly with light, whereas electrically charged particles can influence one another through strong coulombic forces. Hence the dependence on electro-optic devices, which convert light to an electrical signal that alters the optical properties of a material through the coulombic force. The amplification of signals, a necessary part of sequential logic systems, is also difficult to imple-

IMAGE  
UNAVAILABLE  
FOR COPYRIGHT  
REASONS

Another fine mess — the latest in optical computers.

ment entirely optically because of the weakness of interactions. In the systems discussed here, gain is provided by means of electro-optic switches that replace a degraded signal with one fresh from the laser light source.

The builders of the new machine

\* Optical Society of America, Spatial Light Modulators and Applications, Palm Springs, California, 15–17 March 1993.

had modest objectives, proving the principles of operation with components that are well established and available. The computer uses a respectable 66 switches and 200 fibres. This number of components and the stored program mean that it is significantly more complex than previous collections of optical components. Access to the information circulating in the delay line is serial, as in a rotating disk file, and the 20-microsecond memory time is reminiscent of the electronic memories of the 1960s.

Can this be extended to larger, more powerful versions? All physical parameters of the computer will need to be changed by several orders of magnitude for it to make a competitive processor. Simply multiplying the number of components by some factor would not seem advantageous, as the dependence on time of flight would slow a machine in proportion to its physical dimension. And an optical-fibre/lithium-niobate-based system does not seem as susceptible to miniaturization as semiconductor devices. The use of larger numbers of gates would make it impractical to synchronize the elements through the adjustment of each separate connection length. The machine stores 64 16-bit words, 10,000 times fewer than the simplest personal computer. Large increases in the amount of data in memory would make random access most desirable. A more aggressive version of the system based on integrated optics suggests itself, but is far beyond the capabilities of known technology. Propagation velocities and path lengths could not be as well-controlled, nor would lengths be accessible for adjustment.

These accomplishments and others require much ingenious invention and mastery of a great deal of technology to circumvent the physics that obstructs the introduction of optics into the essential logic of a computer. There is no doubt that optical communication will have a role in large computing systems; indeed, products that link major parts of a large system, for example, a disk file to a processing unit, by optical communication are already commercially available. The question is: how far will optical signalling penetrate into the parts that implement the logical action of the machine? The marketplace does not honour technology just because it is clever. Will optical communication between devices or chips, ever offer advantages, rather than merely be possible? And will the low-cost, low-power and high reliability that are needed to build a system of thousands to millions of devices be achieved? □

*Robert W. Keyes is at the IBM T. J. Watson Research Center, Yorktown Heights, New York 10598, USA.*

## Ancient sea water

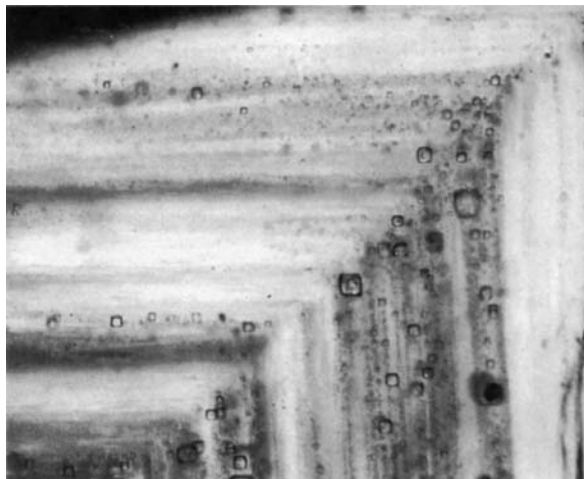
*Paul Knauth*

ANYONE who has ever swallowed a mouthful of sea water has probably wondered why the ocean is salty. The question has never been fully answered, but the most tightly reasoned arguments suggest that chloride was outgassed from primitive rocks with water vapour shortly after the origin of the Earth and that sodium and other principal cations were released from the weathering and hydro-

oceans retreated from the continents. Samples with excessive salinity were explained as sea water that had evolved chemically by reaction with the host rocks. The subsurface brines appeared to be a form of fossil sea water, albeit strongly modified, and their study offered a possible approach for inferring ancient ocean chemistry.

In 1966, the hydrogen and oxygen isotope composition of water in sedimentary basins was convincingly interpreted as indicating that local rainfall had pervasively flushed out all traces of any ancient sea water that might have been deposited with the rocks<sup>3</sup>. At the same time, other isotope studies seemed to indicate that the original mineralogy and geochemistry of the host rocks themselves were obliterated and overprinted by this continuous flux of fresh ground waters that slowly evolved into concentrated brines by water/rock interactions. This one-two punch effectively wiped out hope of finding fossil sea water or of inferring much or anything about the chemistry of the ancient oceans from analyses of sedimentary rocks.

More recently, a rapidly growing number of investigators has begun to reinterpret the chemical and isotope data for both deep ground water and rocks. The isotope data for many subsurface brines can now be reasonably interpreted in terms of mixing of later fresh water with ocean water that experienced various degrees of evaporation on ancient salt flats<sup>4</sup>, and an elegant method has been established for back-tracking chemical evolution of brines to their original seawater chemistry<sup>5</sup>. Numerous workers are presenting strong arguments that marine minerals formed early and can be found in very ancient sediments; the isotope and geochemical record of marine sedimentation has not been obliterated by later alteration<sup>6</sup>. In addition, fluid inclusions trapped along the growth faces of ancient halite crystals are now interpreted by some as evaporation-concentrated sea water that has been preserved for hundreds of millions of years (see micrograph above). Chemical analyses of these inclusions have been used to infer aspects of the isotope and



Fluid inclusions of evaporation-concentrated sea water on growth surfaces of 400-million year-old (Silurian) halite crystal from Michigan basin. Cubic inclusions are negative crystals of the cubic form of halite. Previous efforts to deduce the chemical and isotopic history of the oceans have focused on these types of inclusions, altered by the effects of evaporation.

thermal alteration of rocks. Careful examination of the sequence of evaporite minerals that formed in places where shallow seas once dried up suggests that modern ocean chemistry was achieved at least several hundred million years ago and has remained fairly constant ever since<sup>1</sup>. However, uncertainties in this approach allow plenty of room for significant variations over time, and changes in seawater chemistry are often invoked to explain biological extinctions and secular changes in the mineralogy of ancient sedimentary rocks.

What is needed to evaluate the chemical history of sea water are actual examples of fossil water that can be analysed chemically. Do they exist? Johnson and Goldstein present candidate samples on page 335 of this issue<sup>2</sup>.

The idea that remnants of ancient ocean water are preserved is not new. Ground water extracted from wells in oil fields becomes progressively saltier with depth and eventually reaches the salinity of sea water. Until 1966, it was commonly thought that this salt water was ancient sea water that had seeped down and was left stranded whenever the



chemical history of the hydrosphere<sup>7,8</sup>.

Amidst this ongoing controversy over the origin of subsurface fluids and the preservation of the ancient rock record, Johnson and Goldstein identify tiny mineral cements bearing strong textural, geochemical and isotope evidence that the crystals grew in sea water 500 million years ago and have not been altered since. Extraordinary as that may be, they further observe that the cements contain tiny fluid inclusions. They present evidence that these are small samples of normal sea water that were locked into the growing crystals and have been preserved over geological time — the world's oldest fossil ocean water.

Alternative explanations may exist for their data, but Johnson and Goldstein have assembled several lines of evidence and chosen the simplest explanation. If samples of normal ocean water really are preserved in ancient rocks, a way is open to deduce the chemical history of the oceans by direct measurement. Chemical analysis of such tiny samples — each

contains barely  $10^{-11}$  millilitres (W. J. Johnson, personal communication) — will be difficult, and there is a strong chance that many of the inclusions have been modified during their long entombment. The eventual outcome may yield only a taste of ocean history, but it is bound to be more satisfying than the one that comes with that inadvertent gulp of sea water. □

Paul Knauth is in the Department of Geology, Arizona State University, Tempe, Arizona 85287, USA.

1. Holland, H. D. *The Chemical Evolution of the Atmosphere and Oceans* (Princeton Univ. Press, 1984).
2. Johnson, W. J. & Goldstein, R. H. *Nature* **362**, 335–337 (1993).
3. Clayton, R. et al. *J. geophys. Res.* **71**, 3869–3882 (1966).
4. Knauth, L. P. & Beeunas, M. *Geochim. cosmochim. Acta* **50**, 419–433 (1986).
5. Carpenter, A. B. *Oklahoma geol. Surv. Circ.* **79**, 60–77 (1978).
6. Clauer, N. & Chaudhuri, S. (eds) *Isotopic Signatures and Sedimentary Records* (Springer, Berlin, 1992).
7. Knauth, L. P. & Roberts, S. K. *Stable Isotope Geochemistry. A Tribute to Samuel Epstein* (eds Taylor, H. P. & O'Neill, J. R.) 91–104 (Geochemical Society Spec. Publ., 1991).
8. Das, N., Horita, J. & Holland, H. D. *Geochim. cosmochim. Acta* **54**, 319–327 (1990).

## DEVELOPMENTAL NEUROBIOLOGY

# Recitative and aria

Greg Lemke

TROPHIC factors are powerful regulators of cell proliferation, differentiation and interaction. Now, two long-sought trophic factors of the nervous system — one that triggers cell division in glial cells and a second that potentiates expression of acetylcholine receptors in muscle — have been purified, sequenced and cloned<sup>1,2</sup>. Remarkably, these two factors turn out to be products of one and the same gene, which also encodes putative ligands for the neu (erbB2) receptor tyrosine kinase.

On page 312 of this issue, Marchionni and colleagues<sup>1</sup> report their isolation of complementary DNA and genomic clones encoding the glial growth factors (GGFs), a set of closely related proteins originally defined and purified on the basis of their ability to stimulate DNA synthesis and cell division in cultured Schwann cells<sup>3</sup>. Schwann cells are the principal non-neuronal cells of the vertebrate peripheral nervous system. They participate in a series of complex reciprocal interactions with differentiating neurons, and play essential roles in the development, function and regeneration of peripheral nerves<sup>4</sup>. Although they are highly proliferative during late embryogenesis and in response to nerve injury, Schwann cells are quiescent in mature nerves, or when cultured *in vitro* in the absence of neurons. In the late 1970s, Raff *et al.* identified an apparent-

ly novel activity, present in extracts of bovine brain and pituitary, which strongly stimulated proliferation of these cells<sup>5</sup>.

From this beginning, the journey to the complete purification and cloning of the GGFs has been an arduous one, and has involved the homogenization of many tens of thousands of bovine pituitaries. Several of the basic chromatographic procedures used in the purification of GGF were developed in the early 1980s by Brockes and myself; we gave the factor its name and reported it to be a novel, basic polypeptide with a molecular weight of about 31K (ref. 3). It has taken the better part of ten years, and the methodological refinements of Stroobant, Goodearl and their colleagues at the Ludwig Institute in London, to arrive now at not one, but three major GGF isoforms, designated GGF-I, -II and -III, of 34K, 59K and 45K, respectively<sup>1</sup>. Marchionni *et al.* sequenced tryptic fragments of these proteins, identified a peptide sequence common to all, and then reverse-translated this sequence to generate oligonucleotide probes for the screening of bovine genomic and bovine and human cDNA libraries. These efforts have suggested that GGF-I, -II and -III may be part of an even larger array of as many as eight GGF-related polypeptides, some soluble and some membrane-bound, and all generated by alternative splicing of

the single GGF gene<sup>1</sup>.

But this gene is not a new one. Starting with an assay based on stimulation of tyrosine autophosphorylation of the neu (erbB2) receptor protein-tyrosine kinase, Holmes *et al.* and Wen *et al.* recently demonstrated that this same gene encodes putative neu ligands<sup>6,7</sup>. The 185K neu protein has been intensively studied, as it is closely related in structure to the well-known receptor for epidermal growth factor, and is subject to a point mutation in its transmembrane domain that confers transforming activity. Nonetheless, the identification and purification of its ligands have proved (now not surprisingly) to be nearly as difficult as the purification of the GGFs. In humans these ligands are referred to as heregulins<sup>6</sup>, and in rodents as neu differentiation factors (NDFs)<sup>7</sup>. They seem to be encoded by alternatively spliced transcripts of the same heregulin/NDF/GGF gene, and most (but not all) of the alternatively spliced GGF transcripts described as cDNAs by Marchionni *et al.* have previously been described as heregulin or NDF isoforms. Although they were originally identified because of their ability to stimulate neu autophosphorylation, it remains uncertain whether the NDFs are true ligands for neu, as opposed to a closely related receptor, such as the recently described HER4 protein<sup>8</sup>. Nonetheless, it is clear that the NDFs and the GGFs are products of the same gene.

The second new paper<sup>2</sup> illustrates that the story of the same proteins with several functions and names does not end with the NDFs and the GGFs. Nearly coincident with the long course of work on GGF has been the seemingly unrelated analysis, by Fischbach and colleagues, of a trophic factor that stimulates muscle cell (myotube) differentiation, as assessed by expression of the receptors for the neurotransmitter acetylcholine (AChRs). Designated ARIA, for acetylcholine-receptor-inducing activity, this factor was also first identified in the late 1970s, as a trophic, AChR-inducing activity present in extracts of chick brain and spinal cord<sup>9</sup>. Just as neuronal axons control Schwann-cell differentiation and phenotype, so motor neurons and their presynaptic terminals at the neuromuscular junction dramatically influence the phenotype of developing muscles<sup>10</sup>.

Using an independent purification protocol, Falls and colleagues have now succeeded in purifying ARIA as a 42K protein<sup>2</sup> from chick brain. As in the GGF study, amino-acid sequencing of tryptic fragments of this purified protein was followed by the synthesis of corresponding oligonucleotides, which were in turn used to isolate a series of ARIA cDNAs from an embryonic chick brain

library. The initial ARIA peptide sequence suggested, and the cDNA sequence indeed confirmed, that ARIA is also a product of the NDF/hereregulin/GGF gene<sup>2</sup>. Both purified GGF<sup>1</sup> and purified ARIA<sup>2,11</sup> stimulate tyrosine phosphorylation of a 185K protein, which may be neu or HER4.

*In situ* hybridization experiments performed by both Marchionni *et al.* and Falls *et al.* indicate the NDF/GGF/ARIA messenger RNAs are products of neurons, that expression is particularly high in motor neurons of the embryonic spinal cord, and that the mRNAs are expressed at relatively early stages of embryonic neural development. These expression patterns might be expected for a protein that mediates the stimulatory effects of axons on Schwann cell proliferation and gene expression; they are demanded of a protein that mediates the observed trophic effects of motor neurons and their terminals on muscle development; and they are consistent with the hypothesis that released GGF accounts for the nerve dependence of amphibian limb regeneration<sup>12</sup>. So at least some of the interactions between neurons and their targets seem to be mediated directly by exogenous agents released or presented by neurons, rather than through indirect regulation of the expression of endogenous target cell trophic factors.

The great diversity of biological responses to NDF/GGF/ARIA is not unusual. Indeed, a broad spectrum of activities now seems to be the rule, rather than the exception, for several trophic factors, and it is a real possibility that the isoforms generated by the NDF/GGF/ARIA gene have distinct functions and activities. Sorting out which activity is associated with which protein variant, as well as how the various isoforms may differentially bind to neu or related receptors, is certain to provide still further insights into the biological properties of this expanding trophic-factor family. □

Greg Lemke is at the Molecular Neurobiology Laboratory, The Salk Institute for Biological Studies, 10010 North Torrey Pines Road, La Jolla, California 92037, USA.

1. Marchionni, M. A. *et al.* *Nature* **362**, 312–318 (1993).
2. Falls, D. L., Rosen, K. M., Corfas, G., Lane, W. S. & Fischbach, G. D. *Cell* **72**, 801–815 (1993).
3. Lemke, G. E. & Brookes, J. P. *J. Neurosci.* **4**, 75 (1984).
4. Bunge, R. P. & Griffin, J. W. in *Diseases of the Nervous System. Clinical Neurobiology* 87–100 (W. B. Saunders, Philadelphia, 1992).
5. Raff, M. C. *et al.* *Cell* **15**, 813–822 (1978).
6. Holmes, W. F. *et al.* *Science* **256**, 1205–1210 (1992).
7. Wen, D. *et al.* *Cell* **69**, 559–572 (1992).
8. Plowman, G. D. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **90**, 1746–1750 (1993).
9. Jessell, T. M., Siegel, R. E. & Fischbach, G. D. *Proc. natn. Acad. Sci. U.S.A.* **76**, 5397–5401 (1979).
10. Hall, Z. W. & Sanes, J. R. *Cell* **72**, *Neuron* (Suppl.) 99–121 (1993).
11. Corfas, G., Falls, D. L. & Fischbach, G. D. *Proc. natn. Acad. Sci. U.S.A.* **90**, 1624–1628 (1993).
12. Brookes, J. P. & Kintner, C. R. *Cell* **45**, 301 (1986).

## Where has HIV been hiding?

Howard M. Temin and Dani P. Bolognesi

ONE major question about HIV disease has been “Where is the virus?”. Early in the epidemic it was difficult to find infectious HIV-1, even in people suffering from AIDS, but as techniques have improved it has been possible to demonstrate the virus in more and more infected people. Now papers on pages 355 and 359 of this issue<sup>1,2</sup> show that large amounts of virus are present even during the asymptomatic stage of the disease, although it is not always in the form that might have been expected.

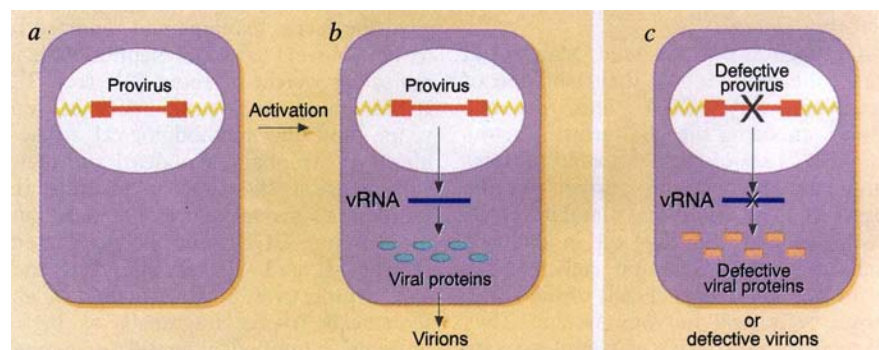
The key to these latest developments is the polymerase chain reaction (PCR), which can measure a mere 1–10 HIV-infected cells present in a total of 100,000 uninfected cells. Pantaleo *et al.*<sup>1</sup> use PCR to show that in asymptomatic HIV-infected patients there are many infected cells in the lymph nodes, more than in the blood, and that these lymph-node cells are more likely to be virus-productive than cells in the blood. Embretson *et al.*<sup>2</sup> extend their earlier work combining PCR and *in situ* hybridization with autoradiography<sup>3</sup> to demonstrate that a high fraction of cells, around a quarter, in germinal centres of lymph nodes are infected, and that these cells tend not to be expressing viral RNA. In some sense these cells with viral DNA and no viral RNA can be called latently infected (see figure).

At first sight, the reader might have difficulty in rationalizing these results. Pantaleo *et al.* stress the point that there is much virus expression in secondary lymphoid organs, whereas Embretson *et al.* tell us that the overwhelming infection in lymph nodes is of a latent nature.

Is there an inconsistency here? We believe not. Pantaleo *et al.* did not attempt to determine the fraction of cells that were latently infected versus those expressing RNA at a quantitative single cell level, as Embretson *et al.* did. Indeed, similar measurements of cells of the blood would render these analyses more complete. The issue of the percentage of latently infected cells notwithstanding, one can conclude from both reports that secondary lymphoid organs are solidly infected with HIV. Both also point out that virus particles are carried, probably passively, by follicular dendritic cells but in a form that is likely to be infectious for other HIV target cells, especially T-cells, that circulate through these organs.

Many questions arise from these data. One in particular is the proportion of cells with viral DNA and no viral RNA that carry infectious as opposed to defective proviruses (see figure). The high rate of errors in retrovirus replication could result in many defective proviruses<sup>4</sup>. If so, cells infected by defective proviruses might not be killed by virus replication nor be subject to host immune surveillance. They might also be protected from infection by non-defective viruses.

It is now clear that there are many cells harbouring virus in an HIV-infected individual, even early after infection, and that the infection takes many different forms including one in which the virus is completely latent. One implication is that, for a preventive vaccine to be effective against HIV-1, the vaccine must block the virus from ever estab-



The differences between a latently infected (a) and a virus-producing cell (b), and between an infectious (b) and a defective (c) provirus. In all of the cells there is an HIV-1 provirus, depicted as integrated HIV-1 DNA in the cell nucleus. In the latently infected cell, the provirus is not transcribed. Thus, there is no viral RNA nor protein. The latently infected cell can be activated to become virus-producing; after activation, early and then late transcription starts, producing viral proteins and ultimately infectious virus. By contrast in the cell infected with a defective provirus, there is a defect (deletion or lethal mutation) in the provirus such that it either makes no viral RNA or viral protein or, as shown, makes defective viral proteins or virions. (There also could be a defective latently infected cell. On activation it would act like the cell infected with a defective provirus.)



lishing residence in the host, particularly the seeding of lymphoid organs. Another is that a curative regimen for HIV-1 infection is likely to be very difficult to develop. Many of the cells in an HIV-1 infected person are latently infected, so they will be resistant to any usual drug or immunological treatments. Even regimens designed to prevent emergence of drug-resistant virus mutants<sup>5</sup> could not overcome such a large reservoir of latently infected cells because inhibitors of reverse transcriptase would not affect activated virus production from latently infected cells. As in many other diseases, the best opportunity for tackling HIV-1 disease would appear to be through prevention and early treatment, preferably upon diagnosis of the initial infection when this is possible.

These reports also indicate that a good start for a more centralized AIDS research effort would be to examine cohorts of infected persons and carefully quantify, in different tissues at different times after infection, cells containing infectious provirus, cells containing defective provirus, cells containing early viral RNA, cells containing late viral RNA, and the load of free virus. From these data, the following can be calculated. (1) What fraction of cells are blocked in early stages of infection (that is, contain only unintegrated viral DNA and no viral RNA<sup>6,7</sup>). (2) What fraction of cells are latently infected (contain viral DNA and no viral RNA<sup>2</sup>). (3) What fraction of latently infected cells have the potential to become productively infected. (4) What fraction of cells are blocked at early stages of infection (contain early viral RNA and not late viral RNA<sup>8</sup>). And (5) what fraction of cells are producing virus (contain late viral RNA). One would also want to know about the dynamics of these stages in an infected individual, and to what extent the dynamics control the variable period between infection and disease.

A more difficult matter is describing the cycling between latently infected cells and infectious virus. Although it would be labour intensive, PCR of proviruses and of RNA from free virions over time in infected persons, followed by sequencing of variable regions, will provide relevant data. Because HIV-1 is so variable as a result of high rates of neutral and selected evolution<sup>9</sup>, the

accumulation of genetic changes can be used to distinguish ancestor and descendant viruses. Another task would be to determine what external signals control the transitions between these stages and how to affect them.

Two final points. First, together with emerging data from other laboratories, these papers show that we have vastly underestimated the extent of virus activity in an infected person, particularly during the 'asymptomatic' phase. This point is emphasized greatly by a study by Piatak *et al.*<sup>10</sup>, who measured unexpectedly high levels of HIV-1 virions in plasma during all stages of infection. To explain the pathogenesis of AIDS many theories have been invoked that have attempted to bypass direct effects of the virus — these results demonstrate that the extent of HIV-1 infection is very significant even during the early stages

after infection, leaving little doubt that the virus itself is sufficient to bring about the disease.

Second, some politicians and activists seem to think that AIDS is a simple problem and that — unlike Jonas Salk and Albert Sabin with polio — scientists are just not good enough to understand it. That is not so. A clear message from the new work is that HIV infection is a very complicated process and that it will require the march of science, frustratingly slow as it sometimes seems, to gain sufficient knowledge to control it. □

Howard M. Temin is in the McArdle Laboratory for Cancer Research, University of Wisconsin—Madison, Madison, Wisconsin 53706, USA. Dani P. Bolognesi is in the Department of Surgery, Duke University Medical Center, Durham, North Carolina 27710, USA.

## PROTEIN ENGINEERING

# Making of the minibody

F. A. Quiocho

THE day when proteins can be routinely designed to have specific functions comes a little closer with the appearance of the paper by Pessi and colleagues on page 367 of this issue<sup>1</sup>. This type of protein engineering has great promise as a powerful research and technological tool in many areas of biology and medicine.

The initial step to be taken when designing a protein is to decide whether to start from scratch or to take advantage of a naturally occurring example and modify it. The first approach has had limited success, because our understanding of protein folding is still fairly rudimentary and so lends another level of complexity to the problem. In contrast, the option of recruiting known protein folds as scaffolds or frameworks for the insertion of functional sites, or the modification of existing enzymatic activity, has proved to be viable and has attracted the most attention.

Pessi *et al.* describe the construction of a protein with a novel  $\beta$ -fold, and they thereby expand the repertoire of protein motifs for grafting 'foreign' functional or novel sites. This protein constitutes the latest (and a highly imaginative) way of producing such a scaffold, and the authors dub it a 'minibody', as the design is based wholly on the structure of part of a variable domain of an immunoglobulin. The 61-residue minibody protein, made by solid-phase peptide synthesis, incorporates three  $\beta$ -strands from each of the two  $\beta$ -sheets of the variable heavy domain ( $V_H$ ) of the antibody, along with the segments corresponding to the ex-

posed hypervariable H1 and H2 loops of the immunoglobulin. The authors judiciously carried out additional residue modifications to make the loop conformations as similar as possible to those observed in the antibody structure. The  $\beta$ -sheet structure serves as a stable platform for the hypervariable loops, the centre for engineering functional sites or displaying epitope libraries.

The behaviour of the minibody in circular dichroic measurements and in denaturation experiments is fully consistent with the expected compact, all- $\beta$  conformation. Further evidence of that conformation was achieved by successfully engineering a zinc-binding site into the minibody with the introduction of three histidine residues in suitable positions, one in H1 and two in H2, that nearly display the geometric requirements and metal selectivities and affinities observed in carbonic anhydrase. The low solubility of the minibody has, however, precluded structure determination by NMR techniques.

This first example of the design and synthesis of an all- $\beta$  protein is a major feat, because producing such a structure is inherently more difficult than making a helical protein. Metal-binding sites have previously been engineered into other proteins, including those with  $\alpha$ -helical frameworks, the active site of enzymes, and the light-chain-complementarity determining regions in an antigen-binding pocket of a native antibody molecule<sup>2-4</sup>. A far more difficult challenge in protein engineering will be to endow the metal-binding site with a

1. Pantaleo, G. *et al.* *Nature* **362**, 355–358 (1993).
2. Embretson, J. *et al.* *Nature* **362**, 359–362 (1993).
3. Embretson, J. E. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **90**, 357–361 (1993).
4. Li, Y. *et al.* *J. Virol.* **65**, 3973–3985 (1991).
5. Chow, Y.-K. *et al.* *Nature* **361**, 650–653 (1993).
6. Zack, J. A. *et al.* *Cell* **61**, 213–222 (1990).
7. Bukrinsky, M. I., Stanwick, T. L., Dempsey, M. P. & Stevenson, M. *Science* **254**, 423–427 (1991).
8. Seshamma, T., Bagasra, O., Trono, D. & Baltimore, D. *Proc. natn. Acad. Sci. U.S.A.* **89**, 10663–10667 (1992).
9. Balfe, P., Simmonds, P., Ludlam, C. A., Bishop, J. O. & Brown, A. J. L. *J. Virol.* **64**, 6221–6233 (1990).
10. Piatak, M. Jr *et al.* *Science* **259**, 1749–1784 (1993).

catalytic activity fully characteristic of carbonic anhydrase or with other zinc-metal-requiring activities<sup>5,6</sup>.

Like their counterparts in immunoglobulins, the H1 and H2 loops of the minibody seem easily to accommodate changes in sequence and insertions. This characteristic is likely to be valuable when it becomes necessary to make modifications for specific protein designs. Indeed, the minibody is further being exploited by the authors, in that

they are constructing a phage-displayed, conformationally constrained peptide library by randomizing the sequence of the hypervariable loops. Its ability to display discontinuous structured epitopes is one great advantage of the minibody in epitope library technology<sup>7</sup>. □

F. A. Quijcho is at the Howard Hughes Medical Institute and the Department of Biochemistry, Baylor College of Medicine, Houston, Texas 77030, USA.

1. Pessi, A. *et al.* *Nature* **362**, 367–369 (1993).
2. Arnold, F. H. & Haymore, B. L. *Science* **252**, 1796–1797 (1991).
3. Tainer, J. A., Roberts, V. A. & Getzoff, E. D. *Curr. Opin. Biotechnol.* **2**, 582–591 (1991).
4. Higaki, J. N., Fletterick, R. J. & Craik, C. S. *Trends*

- biochem. Sci.* **17**, 100–104 (1992).
5. Vallee, B. L. & Auld, D. S. *Biochemistry* **29**, 5647–5656 (1990).
6. Wilson, D. K., Rudolph, F. B. & Quijcho, F. A. *Science* **252**, 1278–1284 (1991).
7. Scott, J. K. *Trends biochem. Sci.* **17**, 241–245 (1992).

## TECTONICS

# Going with the flow

Simon Lamb

GEOLOGISTS are continually reminded by the familiar sight of crumpled and broken rock strata that the continents are not always rigid. Geophysicists have even been tempted to treat the continents, at appropriately large length scales, as a viscous fluid. The assumption is that the brittle crust is weak and closely follows the controlling flow in the underlying ductile and strong part of the lithosphere. But does the brittle crust have any influence on the underlying mantle? According to Jackson, Holt and Haines, in a pair of new papers, it probably does (J. A. Jackson, W. E. Holt & A. J. Haines *J. geophys. Res.* **97**, 17657–17684; 1992; W. E. Holt & A. J. Haines *Tectonics* **12**, 1–20; 1993).

The authors attempt to derive the large-scale motions of the crust in two wide and active plate-boundary zones. These can potentially provide one of the clearest pictures we have of the response of continental lithosphere to forces generated at plate margins. The aim is to see through the detailed effects of deformation in the brittle crust to a more general and fundamental pattern of flow.

The technique was developed over ten years ago by Haines (*Geophys. J. R. astr. Soc.* **68**, 203–209; 1982; also see R. I. Walcott *Geophys. J. R. astr. Soc.* **79**, 613–633; 1984), but until recently has been largely ignored. The original problem was to determine the horizontal motions in the New Zealand plate-boundary zone from repeated triangula-

tion surveys during the past 100 years. Only angular measurements were sufficiently accurate to warrant a comparison between surveys. However, as any surveyor knows, as long as the length of one baseline is known, a knowledge of the angles in a network of triangles is sufficient to determine positions relative to the baseline (Fig. 1).

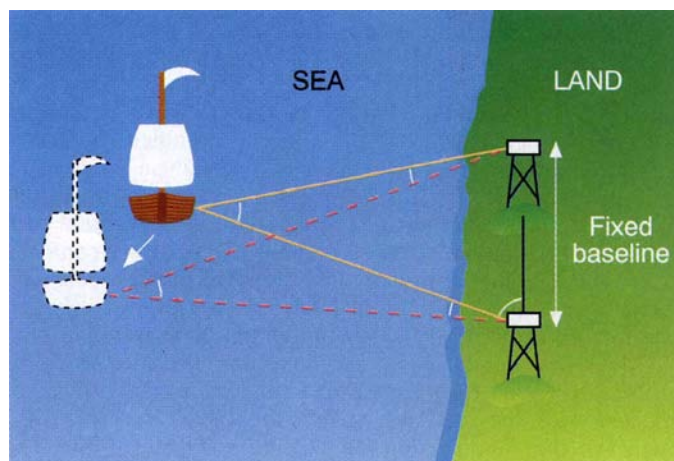


FIG. 1 Tracking the motion of a ship at sea from both measurements at sea and from land-based observatories illustrates the technique of deriving crustal motions from strain measurements in a wide plate-boundary zone. Given the distance between the observatories, the movement of the ship can easily be monitored by triangulation using only angular measurements. The 'land-based' observatories for measuring crustal motions would be on one of the rigid plates.

The key point, exploited by Haines, is that if the baseline does not change length between surveys, it is also possible to calculate displacements merely from the changes in angle since the first survey. These changes in angle can be described in terms of crustal strain, and as no length changes occur on the bounding rigid plates, Haines showed

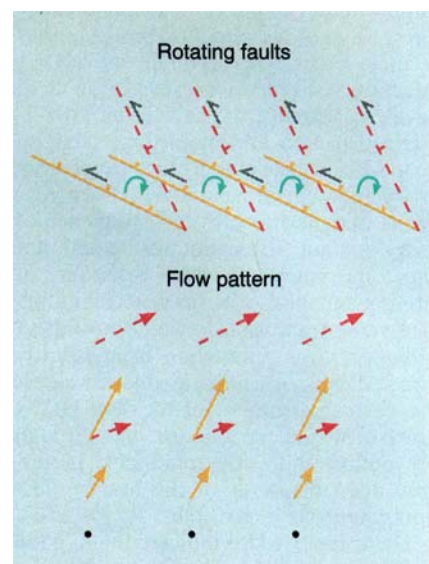


FIG. 2 The flow pattern generated by motion on a set of parallel faults changes as the faults rotate. As slip on the faults cannot result in a length change parallel to the faults, the instantaneous direction of no length change must rotate with the faults.

that the spatial variation in geodetic strain measurements across a plate-boundary zone determines uniquely the displacements within the deforming zone relative to one of the boundaries.

Instead of geodetic measurements, Holt and Haines now use earthquake-moment release to determine local crustal strain in the Aegean (with Jackson) and in Asia. It is straightforward to calculate the average strain, including the absolute change in surface area, accommodated by seismic motion on numerous faults in a region for any particular period. The smoothed distribution of seismic strain throughout the deforming region can be used in the same way as geodetic strain measurements to determine the large-scale pattern of crustal flow. Local changes in surface area act as an important constraint which greatly improves the confidence in the solution. This way the authors find that the direction and magnitude of the motion between

central/eastern Tibet and South China is the same (within 25% error limits) as the relative plate motion between India and Eurasia. The Aegean region is extending rapidly and moving in a southwest direction relative to Europe.

After going to such lengths to 'get away' from the local details of faulting in the brittle crust, it may seem strange that



the authors then invert the flow fields they so derived to determine a pattern of faulting which could have shown them the flow in the first place. The flow field can always be modelled by slip on faults which are parallel to the directions of no length change in the flow. However, because the flow is derived from an averaged and smoothed pattern of crustal strain, these directions need not coincide with the actual pattern of faulting. Yet the synthetic pattern is similar to some of the faulting in both Asia and the Aegean, predicting even the curvature of the major strike-slip faults in eastern Tibet.

The pattern of faults in the brittle crust is only an instantaneous picture; palaeomagnetic measurements suggest that the faults in both Asia and the Aegean have rotated about a vertical axis. If the faults are long-lived features and generally coincide with the instantaneous directions of no length change in the large-scale flow field, then it appears that the latter have also rotated about vertical axes (Fig. 2). This can have happened only if the flow field has changed with time. But as the authors show in both cases, the rotation of the faults can also be predicted from the crustal flow field at any instant. The conclusion is that the instantaneous pattern of dominant faulting also constrains how the large-scale flow field changes with time. Clearly, if the crust is weak, and conforms to a bulk flow determined by the deformation in the underlying part of the lithosphere, this conclusion makes no sense. Jackson, Holt and Haines suggest that the strength of the crust cannot be neglected. It breaks only in certain directions. Faulting in these directions influences the overall deformation of the continental lithosphere.

How robust is this conclusion? I suspect that it is based on the fact that relatively few earthquakes contribute substantially to the calculated flow in both Asia and the Aegean. Seven (out of 74) events contribute nearly 40% of the total scalar moment in the Aegean — a region  $800 \times 500 \text{ km}^2$ . In Asia, a similar number of earthquakes dominate the moment release in an even bigger region,  $2,000 \times 2,000 \text{ km}^2$ . Faulting in these big events may determine the directions of no length change in the derived flow. Can this be a true picture of the active deformation? Field observations indicate a much more dense pattern of active faulting. Unlike geodetic measurements, which give the total motion in any particular period, earthquake-moment release is only one part of the short-term crustal deformation.

Much of the short-term deformation will be accommodated by elastic straining and also possibly by aseismic slip. In Asia and the Aegean the strain

accommodated by seismic faulting this century is less than the elastic limit of the brittle crust ( $10^{-5}$ – $10^{-4}$ ). The earthquakes are probably not telling the whole story. This is not to say that the derived pattern of flow is likely to be grossly unrepresentative, but rather that one should treat the details of the flow pattern with caution.

Perhaps it will not be until we have the results of satellite-based surveying techniques, which are capable of detect-

ing the total motions in wide plate-boundary zones, that we will be able to confidently assess the role of short-term seismicity in recording the deformation of the continental lithosphere. These results will soon be available for the Aegean, but it may be many years before we can study Asia this way. □

Simon Lamb is in the Department of Earth Sciences, University of Oxford, Parks Road, Oxford OX1 3PR, UK.

## MATERIALS SCIENCE

# Cracks take a new turn

M. Marder

If you have had the maddening experience of watching a crack slowly make its way across a car's windscreen, you will know how complicated the path it leaves can be. At unpredictable points, the crack may wiggle, turn an abrupt corner, or even split into two. Now A. Yuse and M. Sano of Tohoku University have devised an experiment that allows such crack motion to occur in a thoroughly

controlled and reproducible setting (*Nature* **362**, 329–331; 1993). The types of pattern they describe in this issue, and the manner in which the patterns occur, are reminiscent of phenomena well known in solidification of metals, or in fluid flow, but they have not previously been measured for fracture. Furthermore, the pattern-forming process seems to provide a detailed measurement of a fundamental quantity in fracture, the fracture energy, which cannot easily be obtained in other ways.

The experiment of Yuse and Sano involves pushing a thin glass plate from a hot region to a cold one, so that a temperature gradient travels across it. A crack is introduced by the experimenters, who notch the plate, and as the plate begins to move the crack tip jumps just ahead of the thermal gradient, and remains there. When the strip moves at low velocities, the crack forms a perfect stable cut through the centre of the strip. As the velocity increases, however, there is a transition, and the fracture path begins to oscillate, producing the remarkable spectacle of a piece of glass slowly and spontaneously being cut in the shape of a sine wave (Fig. 1). Further increases in velocity bring crack branching, and complicated sets of multiple cracks.

A close parallel to Yuse and Sano's experiment is directional solidification (Fig. 1; J. S. Langer *Rev. mod. Phys.* **52**, 1–28;

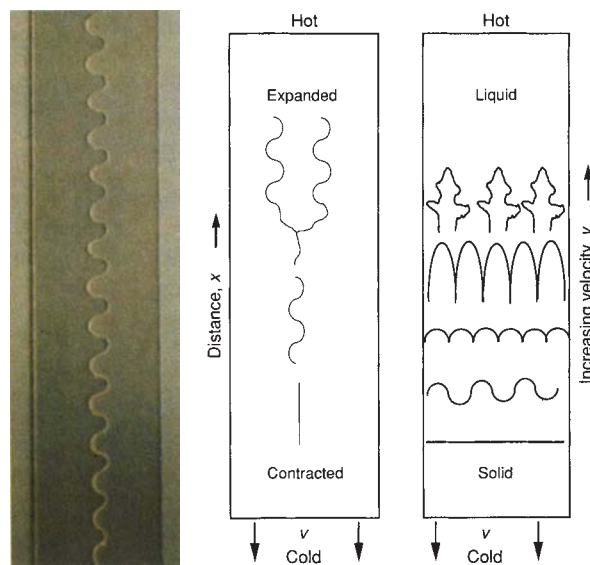


FIG. 1 Fracturing and freezing. Left, photo of sinusoidal cracking obtained by Yuse and Sano. Centre, sketch of the authors' more general results. Right, a sketch of directional solidification. Both experiments feature strips moving from hot to cold regions, and exhibit a sequence of instabilities as the sliding velocity increases. In each drawing, successive figures in the vertical direction represent steady-state patterns achieved in successive experiments where the constant velocity  $v$  increases. In the case of directional solidification, patterns form at the interface between a hot melted region at the top and a cold solid region below. For slow sliding velocities  $v$ , the solidification front is just a straight line, but at some critical velocity, the front begins to ripple. At higher velocities, the ripples become more pronounced, and eventually transform into fingers, which grow in parallel, leaving deep grooves behind them. At still higher velocities, the fingers also become unstable, and create a complicated branching pattern. Yuse and Sano find a similar trend from simple to complicated patterns.

## Reaction cycles

MOBILE chemical patterns are familiar from self-catalysing reactions like the Belousov–Zhabotinsky, but S. L. Lane and D. Luss have found a conventional catalytic process that literally runs rings around itself (*Phys. Rev. Lett.* **70**, 830–832; 1993). The process involves the oxidation of hydrogen gas over a nickel catalyst, shaped as a ring 3 cm across. The authors found that the ring did not remain uniformly hot. One side (about a third of the circumference) was always hotter than the rest; moreover, this hot arc slowly rotated, taking 9½ minutes to complete one cycle. It changed shape and speed during a cycle, but the pattern always repeated itself each time round. Lane and Luss watched the process for over 70 hours. The effect is related to the heat given out by the oxidation and the catalyst's sensitivity to temperature. The authors suggest that even more complicated patterns may arise on commercial catalysts.

## Intolerant mice

An intolerant disposition can be pushed over the limit by no more than a 'dirty' life, according to data on a new mouse model for the autoimmune disease multiple sclerosis (MS) reported by J. Goverman *et al.* (*Cell* **72**, 551–560; 1993). Transgenic mice with a majority of T cells bearing receptors for the nerve sheath protein myelin basic protein (MBP), a possible autoantigen target in MS, provide large numbers of cells of single specificity amenable for studying the breakdown of tolerance in the immune system. These mice are more susceptible than the parental strain to autoimmune disease triggered by immunization with MBP and pertussis, and also, surprisingly, by life in a non-sterile, microbe-filled, environment.

## Cool work

THE most fragile possible molecule has been made by chemists at the University of Minnesota (F. Luo *et al.* *J. chem. Phys.* **98**, 3564–3567; 1993). It comprises a pair of helium atoms. Although helium condenses to make a liquid at 4.2 K, and small clusters of helium atoms have been made before, the diatomic molecule barely survives at 1 mK (the N<sub>2</sub> bond is 100,000,000 times stronger). No classical synthesis was involved in the preparation, merely cooling helium gas by squeezing it through a pinhole into a vacuum. The molecules were identified mass spectrometrically. The potential well holding the atoms together is so shallow that any vibration breaks the bond. The helium interaction, involving the weakest of van der Waals forces, has been a test of quantum theory from 1928, since when prediction of its strength has been the subject of over 50 papers.

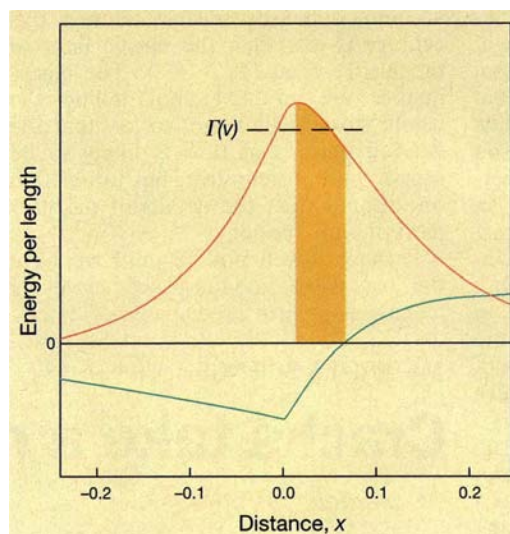


FIG. 2 The energy available per unit length for crack propagation in the experiment of Yuse and Sano (red curve). The horizontal axis shows the position of the crack tip relative to the cold end of the temperature gradient, in units of the width of the plate. The scale of the vertical axis is not indicated, but the only important fact is that the vertical scale is proportional to the square of the temperature difference  $\Delta T$  along the strip. The crack will propagate with its tip at a position given by the intersection of the red curve with the dashed line labelled  $\Gamma(v)$ . The green line monitors the stress field responsible for the stability of the crack. When the crack tip reaches the position where this quantity is positive, it becomes unstable; the orange region indicates the range of stable crack positions.

1980). In directional solidification, a molten solid is pushed at a constant speed into a cold bath, where it hardens. When casting metals, it might seem desirable to push as rapidly as possible, but instabilities intervene. As the pushing velocity increases, the interface between solid and liquid undergoes a series of transitions, forming progressively more and more complicated patterns. These structures alter the mechanical properties of metals, and provide a fascinating example of how natural phenomena can create complicated patterns from almost completely homogeneous external conditions.

The reason that cracks propagate in Yuse and Sano's experiment can be understood by thinking about the effect of thermal gradients upon the glass plate. In the hot region, the plate is slightly expanded, while in the cold region it is slightly contracted. Since the hot and cold regions are glued seamlessly together, there must be stresses where they join. The hot thick portion of the strip is pinched where it joins the cold narrow portion, and its effort to spring back from the pinching provides the forces needed for crack propagation. One can think of a crack as an object that sucks stress energy out of a plate in the same way a vacuum cleaner picks up dust. At any given speed  $v$ , the crack

requires a certain amount of energy per unit length  $\Gamma(v)$  to move, so that if more energy is provided by the plate, the crack will use it to accelerate.

The red curve in Fig. 2 shows the energy per length that the strip provides to a straight crack running along its centre. The position of the crack is measured with respect to the point at which the strip is cooled; far behind this point, or far ahead of it, the energy available for crack motion vanishes. The height of the curve is proportional to the square of the temperature difference  $\Delta T$  along the strip. At any given speed  $v$ , the tip of the crack will find a spatial location where the energy that it derives from this curve equals the energy it needs,  $\Gamma(v)$ , a location always in front of the temperature gradient, for precisely the same reason that a surfer travels in front of a wave rather than behind it. If the temperature  $\Delta T$  rises, the red curve grows in scale, pushing the tip of the crack forward.

The striking feature of Yuse and Sano's experiment is not that the cracks move, but that simple straight line motion becomes unstable in a reproducible way. The condition for instability of a

straight crack has been worked out by B. Cotterell and J. R. Rice (*Int. J. Fract.* **16**, 155–169; 1980) and is represented by the green line in Fig. 2. When this quantity, a tension parallel to the crack, becomes positive, the slightest imperfections in the plate will be amplified and quickly cause the crack to leave the central axis.

So long as the tip of the crack lies within the range marked by the orange region of Fig. 2, crack motion is stable; but once the temperature difference  $\Delta T$  becomes too large, the crack tip is pushed forward of it, and oscillations begin. The new experiment can even be used as a tool to measure in detail the velocity dependence of  $\Gamma(v)$ :  $\Gamma(v)$  is very nearly proportional to the square root of the temperature difference at which the instability first shows for a given tip growth rate,  $v$ .

Not only do these lovely experiments bring fracture within the family of problems that may be studied by techniques of pattern formation and nonlinear dynamics, they should also teach new lessons about the processes of fracture. Maybe it will be possible to stop them ruining car windscreens. □

M. Marder is in the Department of Physics, University of Texas, Austin, Texas 78712, USA.



# Bridging the gap

Graham Warren

A CLEAR view of the operation of the membrane fusion machinery inside cells emerges with the publication of two remarkable papers on pages 318 and 353 of this issue<sup>1,2</sup>, both from James Rothman and colleagues. They point to fundamental similarities in the way all membrane vesicles dock with and thereby deliver their contents to the right membrane compartment.

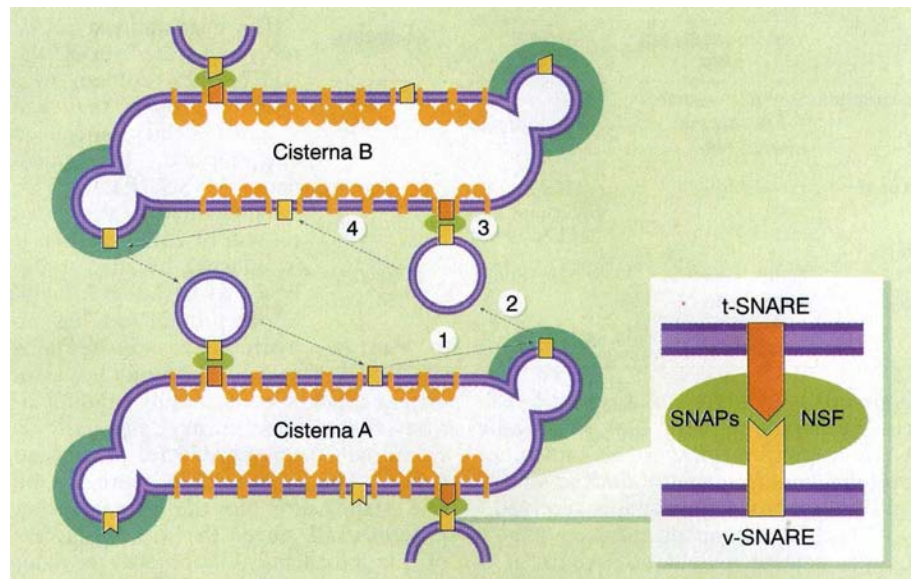
Vesicles deliver selected proteins and lipids to most of the membrane compartments in a eukaryotic cell. They range from transport vesicles carrying newly synthesized proteins from cisterna to cisterna in the Golgi stack<sup>3</sup>, to synaptic vesicles carrying neurotransmitters for delivery to presynaptic membranes<sup>4</sup>. For delivery to occur there must be a docking mechanism that determines precisely which membrane a vesicle is to bind to and a machine to fuse the two lipid bilayers. Rothman and colleagues had earlier characterized several components of a fusion machine which has been shown to operate at many steps on the secretory and endocytic pathways. They now provide good evidence that it also operates at regulated secretory steps such as occur at the nerve terminal. More importantly, they have used the components of this fusion machine to fish out the docking mechanism, exploiting the fact that many of the proteins in synaptosomal membranes have been cloned and sequenced. They not only identify two of these proteins as components of the docking mechanism, but their results also provide a general model that could explain the specificity of vesicle docking<sup>1</sup>.

Using a cell-free assay that faithfully reconstitutes the transport of vesicles between the *cis* and *medial* cisternae of the Golgi stack, Rothman and colleagues had identified a 76K homo-tetramer as a component of the fusion machine. This *N*-ethylmaleimide sensitive factor (NSF) could not bind directly to membranes but needed three SNAPs (soluble, NSF accessory proteins) termed  $\alpha$ ,  $\beta$  and  $\gamma$ .  $\alpha$ -SNAP and  $\beta$ -SNAP appeared to be interchangeable and, as they now report, this is because  $\beta$ -SNAP is a brain-specific isoform that is almost identical to  $\alpha$ -SNAP<sup>2</sup>. The SNAP receptors could be solubilized in detergent and, together with NSF and SNAPs, formed a complex that could be disassembled by the addition of Mg-ATP, an event probably mimicking that which occurs *in vivo*. By attaching this complex, via NSF, to a column and using salt-washed brain membranes as the source of receptors, they could specifically elute SNAP

receptors using Mg-ATP. Only three proteins were eluted, two of which were syntaxin (both the A and B isoforms) and synaptobrevin-2 (also known as VAMP-2).

Synaptobrevin is a small, integral membrane protein exposed on the surface of synaptic vesicles<sup>5,6</sup>. Syntaxins are also integral membrane proteins with most of their mass in the cytoplasm, but they seem to be present on the presynap-

found in tissues other than brain. Perhaps the best evidence comes from the unusual topology of both proteins. They belong to what may be a new class of membrane protein which has no signal sequence, most of its mass (the amino-terminal part) in the cytoplasm, a single, membrane-spanning domain, and at most a few charged amino acids on the luminal side of the membrane<sup>8</sup>. In this class are a large number of yeast proteins required for secretion which show some sequence similarity to either synaptobrevin or syntaxin (see table<sup>5-7,9-13</sup>). Unfortunately, as has happened so often before, no function had been attributed



Proposed cycle of v-SNAREs between Golgi cisternae. A vesicle budding from the dilated cisternal rims will contain one type of v-SNARE (1) which is exposed, ready for docking, after removal of the COP coat (2) (ref. 3). The vesicle then docks (3; see enlargement) with the partner t-SNARE, and this is then followed by fusion which delivers the v-SNARE to the cisterna (4). The v-SNAREs then cycle back to the previous cisterna by the same mechanism. The t-SNAREs in each cisterna are part of large oligomers each containing different Golgi enzymes (orange). There would need to be two types of oligomer in each cisterna to cope with two cycles of v-SNAREs. The oligomers are shown on opposite faces only for convenience and the cisternae have been separated for clarity.

tic membrane<sup>7</sup>. The importance of this result cannot be overstated because the topology and location of these proteins immediately suggest a simple docking mechanism in which NSF/SNAPs link the synaptobrevin in the vesicle to the syntaxin in the presynaptic membrane (see figure). NSF/SNAPs are general components of the fusion machinery, so the specificity must reside in the interaction between what the authors term v-SNAREs (vesicle-SNAP receptors; synaptobrevin) and t-SNAREs (target-SNAP receptors; syntaxin). Different v- and t-SNARE partners could therefore explain the docking specificity at each of the many vesicle-mediated steps in eukaryotic cells.

Evidence for such an idea comes from the different isoforms of both synaptobrevin and syntaxin, and the fact that immunologically related proteins are

to any of these proteins beyond allocating them to a particular step on the secretory pathway, based on a morphological analysis. The system described by Rothman's group will resolve this problem. Not only can the NSF/SNAPs be used to test whether any of the yeast proteins listed in the table are v- or t-SNAREs, but they can be used to search for new SNAREs especially if purified membranes are available.

Docking must be a regulated process, perhaps one monitored by the rab/YPT family of small GTP-binding proteins. Members of a closely related family ensure the fidelity of protein synthesis by acting as molecular clocks<sup>14</sup>. Vesicles will, presumably, sample many membranes before the v-SNARE meets the correct t-SNARE. Recognition will cause the vesicle to linger, giving enough time for the hydrolysis of GTP which

would then lock the vesicle onto the membrane. In order to monitor the specific interaction between the SNAREs, different rab/YPTs would be needed for each pair, which would explain why there is such a large family<sup>15</sup>.

Docking need not be followed immediately by fusion, especially during regulated secretion. Synaptic vesicles are docked at the presynaptic membrane presumably to ensure a rapid response, triggered by an influx of calcium ions. Syntaxin binds to calcium channels<sup>7</sup> and was originally isolated as a protein

teins from being returned to the previous cisterna. Even small differences in, for example, the size of the vesicle would ensure net forward movement, though at the expense of futile cycling. The other interesting consequence of this model is that each cisterna would contain two types of t-SNAREs, each serving an adjacent cisterna. If each t-SNARE defines its own oligomer, then two such oligomeric types would exist in each cisterna and each oligomeric type would be present in adjacent cisternae. This would give rise to an overlapping distribution of Golgi enzymes

as has recently been observed<sup>19</sup>.

The third and last SNAP receptor eluted from the NSF/SNAPs column was another protein that had already been sequenced and termed, by coincidence, SNAP-25 (for synaptosomal associated protein of 25K; ref. 20). It is unusual because it behaves as an integral membrane protein yet has no hydrophobic stretches in its sequence. Membrane association could be conferred by palmitoylation of a cysteine cluster, especially as palmitoyl-CoA is needed for membrane fusion<sup>3</sup>. This protein could, therefore, be one step nearer to the actual fusion process itself. Given the striking success of the biochemical approach to date, there seems little doubt that it will soon lead to a complete description of the docking and fusion mechanisms. □

Graham Warren is in the Cell Biology Laboratory, Imperial Cancer Research Fund, PO Box 123, Lincoln's Inn Fields, London WC2A 3PX, UK.

- Söllner, S. et al. *Nature* **362**, 318–324 (1993).
- Whiteheart, S. W. et al. *Nature* **362**, 353–355 (1993).
- Rothman, J. E. & Orci, L. *Nature* **355**, 409–416 (1992).
- Südhof, T. C. & Jahn, R. *Neuron* **6**, 665–677 (1991).
- Trimble, W. S., Cowan, D. M. & Scheller, R. H. *Proc. natn. Acad. Sci. U.S.A.* **85**, 4538–4542 (1988).
- Baumert, M., Maycox, P. R., Navone, F., de Camilli, P. & Jahn, R. *EMBO J.* **8**, 379–384 (1989).
- Bennett, M., Calakos, N. & Scheller, R. H. *Science* **257**, 255–259 (1992).
- Kutay, U., Hartmann, E. & Rapoport, T. A. *Trends Cell Biol.* **3**, 72–75 (1993).
- Newman, A. P. et al. *Molec. cell. Biol.* **12**, 3663–3664 (1992).
- Shim, J., Newman, A. P. & Ferro-Novick, S. J. *Cell Biol.* **113**, 55–64 (1991).
- Dascher, C., Ossig, R., Gallwitz, D. & Schmitt, H. D. *Molec. cell. Biol.* **11**, 872–885 (1991).
- Gerst, J. E., Rodgers, L., Riggs, M. & Wigler, M. *Proc. natn. Acad. Sci. U.S.A.* **89**, 4338–4342 (1992).
- Hardwick, K. G. & Pelham, H. R. B. *J. Cell Biol.* **119**, 513–521 (1992).
- Thompson, R. C. *Trends biochem. Sci.* **13**, 91–93 (1988).
- Pfeffer, S. R. *Trends Cell Biol.* **2**, 41–46 (1992).
- Brose, N., Petrenko, A. G., Südhof, T. C. & Jahn, R. *Science* **256**, 1021–1025 (1992).
- Machamer, C. E. *Trends Cell Biol.* **1**, 141–144 (1991).
- Nilsson, T., Lucocq, J. M., Mackay, D. & Warren, G. *EMBO J.* **10**, 3567–3575 (1991).
- Nilsson, T. et al. *J. Cell Biol.* **120**, 5–14 (1993).
- Oyler, G. A. et al. *J. Cell Biol.* **109**, 3039 (1989).

## Perfect stress

TERMITES are notorious for weakening a wooden structure till it is just on the point of collapse. Daedalus reckons that they sense the stresses in the wood as they chew it, and preferentially eat the unstressed material. Where the remaining wood is fully stressed and about to break, they stop eating. This is an elegant method of structural optimization. It leaves the lightest structure that can just carry the applied load. Daedalus is now using chemistry to do for metal what termites do for wood.

Chemical corrosion often has a converse, contra-termitic action. It tends to attack the stressed regions first, where the local energy is higher. This is stress-corrosion. But some perverse reagents protect the surface they are attacking. Thus oxygen 'passivates' chromium, and nitric acid can passivate iron, in both cases by covering the surface with an inert coating of oxide. DREADCO chemists are therefore devising a stress-passivator that preferentially attacks stressed metal, and coats it with a protective layer. DREADCO's cunning 'Termite Solution' is simply a mixture of stress-passivator with a corrosive but stress-insensitive acid.

Thus mathematical stress-analysis, the bugbear of so many engineering designers, will be neatly bypassed. Simply take a scale model, or even the full-scale product itself, stress it up with springs and weights to its maximum permitted load, and just dump it in Termite Solution. An understressed region of metal, not earning its structural keep, will be steadily dissolved by the acid. As it shrinks, the applied load will be carried by less and less metal, raising the local stress. When the region is almost at breaking point, the stress-passivator will move in, coating and protecting it from further attack. Thus the whole structure will be automatically sculpted into the lightest and neatest form that can just do the job. Spare and elegant, it will carry no superfluous weight at all.

Architects and engineers, particularly in the weight-conscious aerospace business, should rush to use Termite Solution. Crudely designed airframes, satellite bodies, pressure vessels, rocket motors and so on, will be loaded or pumped up to design stress, dunked in the solution, and left to optimize. In due course the perfected products will be hauled out and hosed down, ready for service. Even existing structures might be optimized in this way. Daedalus is working on a thixotropic Termite Paste to be smeared over old overdesigned bridges, cranes, tanks, engines, and so on. A new, slimmer, more beautiful creation will then emerge from the bulbous old cocoon.

David Jones

| PROTEINS THAT MAY BE INVOLVED IN VESICLE DOCKING <sup>5-7,9-13</sup> |                                         |                                       |                 |
|----------------------------------------------------------------------|-----------------------------------------|---------------------------------------|-----------------|
|                                                                      | Vesicle-mediated step                   | v-SNARE                               | t-SNARE         |
| Mammals                                                              | Synaptic vesicle → Presynaptic membrane | VAMP/ Synaptobrevin                   | Syntaxin        |
| Yeast                                                                | ER → Golgi                              | Bos1p<br>Sly2p/Sec22p<br>Sly12p/Bet1p | Sed5p           |
|                                                                      | Golgi → Plasma membrane                 | Snc1p, Snc2p                          | Sso1p,<br>Sso2p |
|                                                                      | Golgi → Vacuole                         | ?                                     | Pep12p          |

bound to synaptotagmin, a synaptic vesicle membrane protein that specifically binds to calcium ions<sup>16</sup>. This and other proteins could arrest the docked vesicle until the appropriate signal is received.

If each vesicle-mediated step uses a unique pair of v- and t-SNAREs, it is essential that the t-SNAREs are restricted to the correct membrane compartment. Several members of this new type of membrane protein are retained in the correct compartment because of information in the spanning domain (see ref. 8). The same is true for Golgi enzymes<sup>17</sup> and, as noted more recently, some ER proteins. Taking the Golgi stack as an example (see figure), the presence of a particular t-SNARE would determine which vesicles fuse with the cisterna. Delivery of newly synthesized Golgi enzymes that interact with the t-SNARE through the spanning domain would lead to retention in the cisterna by the formation of oligomers too large to enter the vesicles budding from the dilated cisternal rims<sup>18</sup>. The t-SNARE would therefore determine the composition and hence function of the cisterna.

A problem arises because the v-SNAREs are membrane proteins and so must be recovered from the cisterna into which they are delivered by vesicle fusion. They could simply be incorporated into vesicles that return to the previous cisterna and recognize the same t-SNARE (see figure). There would have to be differences between the forward- and backward-moving vesicles simply to prevent forward-moving pro-



# Recurrence of a binding motif?

**SIR** — The structure of histidine-containing phosphocarrier protein (HPr) has been analysed in detail by both X-ray crystallography<sup>1,2</sup> and nuclear magnetic resonance (NMR) spectroscopy<sup>3</sup>, and the determination of the unphosphorylated form<sup>1</sup> led to a proposal for a detailed model for the phosphotransfer cycling mechanism. Given that proteins with related functions often adopt similar topologies and that reversible phosphorylation is frequently observed *in vivo*, it is surprising that no related structures have been discovered. The structures of the RNA binding domain of ribonucleoproteins (RNP) A and C<sup>4,5</sup> and acylphosphatase (APt)<sup>6</sup> are known, and although the structural similarity between the ribonucleoproteins and acylphosphatase was detected, it was concluded that HPr had a different topology because it did not intercalate two  $\beta$ - $\alpha$  units in the same manner<sup>6</sup>.

We have used coordinates from the Protein Data Bank<sup>7</sup> and a flexible structural alignment program<sup>8</sup> to reassess these structures, and find that HPr also adopts this fold. In fact, a superposition of HPr with APt aligns 61 residues with a root-mean-square deviation for C $\alpha$  atoms of only 2.7 Å despite the existence of only three residue identities in the entire alignment. The structurally equivalent regions of HPr and APt consist of three antiparallel  $\beta$ -strands and two  $\alpha$ -helices (see figure).

Previous observations that the inserted helix is not detected in the NMR structure<sup>3</sup> reaffirms the peripheral importance of this region. The ribosomal protein L7/L12 (ref. 9) is also known to adopt a similar fold, but we find that its relationship with HPr is more distant, with an r.m.s. deviation of 4.0 Å over 54 C $\alpha$  atoms. It is interesting that the pattern of insertions and deletions differs from those between HPr and APt. This situation is also observed in other structures, such as nerve growth factor and transforming growth factor  $\beta$ -2, where much larger insertions and deletions are accommodated within the same fold<sup>10</sup>.

HPr, APt and RNP are all involved in phosphate binding. In HPr, phosphate is covalently bound to His 15 and forms an electrostatic interaction with Arg 17. Although APt performs a different reaction, it too must bind phosphate, though in a noncovalent manner. To achieve this it would also require a positively charged moiety such as a guanidinium group. Chemical modifications of arginine residues in APt appear to inactivate the enzyme<sup>6</sup>, and one of these (Arg 23) is located at the amino terminus of the first helix. Not only is this an ideal candidate for phosphate binding (as the positive charge on the arginine will be enhanced by the helix dipole), but our structural alignment shows that Arg 23 of APt is in the same region as the HPr active site. In addition,

Arg 23 is the only arginine remaining conserved over an alignment of 12 homologous sequences<sup>6</sup>. A similar situation is observed in RNP, where mutations of lysine residues to glutamine at the amino terminus of the first helix significantly reduce RNA binding<sup>4</sup>.

Thus, the observed structural similarity, combined with the correspondence of functionally important, positively charged residues at the amino terminus of the first helix, strongly suggests that the observed fold represents a stable phosphate-binding motif.

**Mark B. Swindells**

Protein Engineering Research Institute,  
6-2-3 Furuedai, Suita,  
Osaka 565, Japan

**Christine A. Orengo**

**David T. Jones**

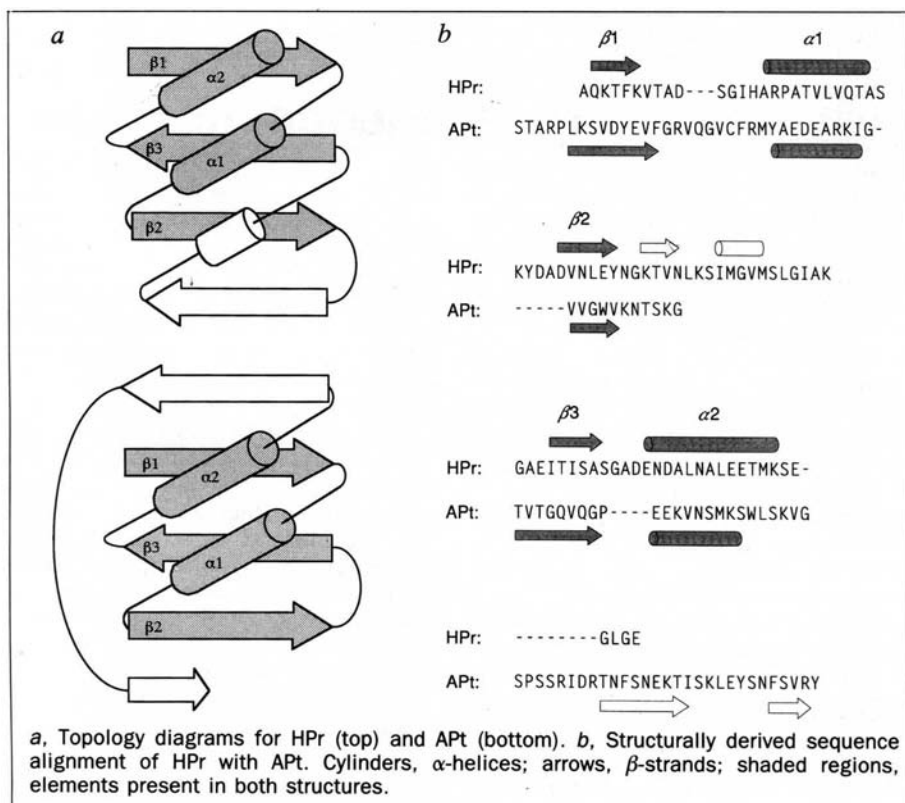
**Laurence H. Pearl**

**Janet M. Thornton**

Biomolecular Structure and Modelling  
Unit,

Department of Biochemistry and  
Molecular Biology,  
University College London,  
Gower Street,  
London WC1E 6BT, UK

1. Zongchao, J., Vondonselaar, M., Quail, J. W. & Delbaere, L. T. J. *Nature* **361**, 94–97 (1993).
2. Herzberg, O. et al. *Proc. natn. Acad. Sci. U.S.A.* **89**, 2499–2503 (1992).
3. Wittekind, M., Reizer, J. & Kleit, R. E. *Biochemistry* **29**, 7191–7200 (1990).
4. Nagai, K., Oubridge, C., Jessen, T. H., Li, J. & Evans, P. R. *Nature* **348**, 515–520 (1990).
5. Wittekind, M., Görlach, M., Friedrichs, M., Dreyfuss, G. & Mueller, L. *Biochemistry* **31**, 6254–6265 (1992).
6. Pastore, A., Saudek, V., Ramponi, G. & Williams, R. J. P. *J. molec. Biol.* **224**, 427–440 (1992).
7. Bernstein, F. C. et al. *J. molec. Biol.* **122**, 535–542 (1977).
8. Orengo, C. A., Brown, N. P. & Taylor, W. R. *Proteins* **14**, 139–167 (1992).
9. Leijonmarck, M. & Liljas, A. *J. molec. Biol.* **195**, 555–580 (1987).
10. Swindells, M. B. *Science* **258**, 1160–1161 (1992).



## Food for thought

**SIR** — Peter Moore (*Nature* **361**, 304; 1993), in his News and Views discussing our paper on apparent convergence between stick-insect eggs and ant-dispersed seeds, wonders how a stick-insect nymph might fare should it hatch from an egg in an active ant nest. S. G. Compton and A. B. Ware (*Psyche* **98**, 207; 1991), whose paper reported similar conclusions to our own, made the interesting observation that stick insect nymphs hatched in a laboratory ant colony were completely ignored by the ants. The lack of chemical recognition of nymphs as potential food items by ants might bear further investigation.

**Lesley Hughes**

**Mark Westoby**

School of Biological Sciences,  
Macquarie University,  
North Ryde,  
New South Wales 2109, Australia

## A purple protist

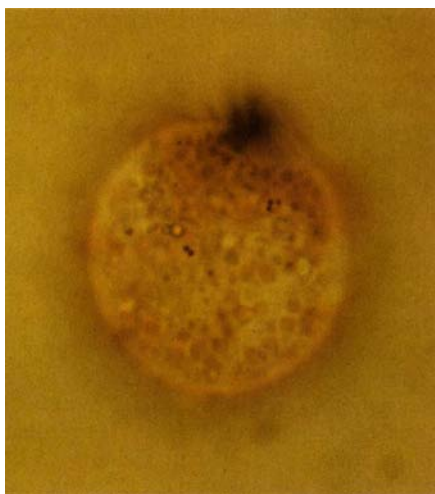
**SIR** — Mitochondria are believed to be descended from endosymbiotic bacteria, which belong to the  $\alpha$ -group of the proteobacteria, because of the similarities of their cytochromes *c* and 16S ribosomal RNA sequences<sup>1</sup>. It is generally speculated that these symbionts were heterotrophic aerobes which established themselves within an anaerobic host, but it is difficult to imagine how such association could be brought about or how it could initially be adaptive to any of the components. Woese<sup>2</sup> has suggested that the symbionts were originally photosynthetic nonsulphur bacteria and that their significance in terms of respiration arose later. Until now, however, no example of endosymbiotic bacteria has been found which could serve as a contemporary model for the origin of mitochondria.

We have recently isolated a ciliate *Strombidium purpureum* Kahl from the photic, upper 5 mm of anaerobic, sulphide-containing marine sands<sup>3</sup> (see figure). We find that the ciliate always harbours 200–700 purple endosymbionts 1.5–2.7  $\mu\text{m}$  in size along the periphery of the host cell. These endosymbionts are not surrounded by vacuoles, but they are covered by a two-layered 11–12-nm-thick cell wall. The symbionts contain photosynthetic membranes in the form of tubules and vesicles arising from invaginations of the cell membrane. The bacteria contain bacteriochlorophyll *a*, the carotenoid spirilloxanthin and storage vacuoles with poly- $\beta$ -hydroxybutyrate, and multiply by budding; these traits suggest a relation to the genus *Rhodospirillum*.

The ciliates are phagotrophs and feed on bacteria. They could be grown under strictly anaerobic conditions in the light. The protozoa show a photosensory behaviour and accumulate at wavelengths corresponding to the absorption spectrum of the symbionts. In the light the cells avoid even traces of oxygen, but in the dark the cells accumulate under microaerobic conditions (1–4% atm. sat.). Under anaerobic conditions the ciliates die within 36–48 hours in the dark, but under microaerobic conditions

they survive darkness 2–3 times longer. Exposed to atmospheric  $p_{\text{O}_2}$  the cells die within 30 min.

We suggest that this protozoon, like many other ciliates<sup>4</sup>, has evolved an anaerobic lifestyle from aerobic ancestors and that the symbiosis with the photosynthetic bacterium evolved in an anaerobic host. Anaerobically and in the light the bacteria utilize the metabolic endproducts of the host (acetate, lactate,  $\text{H}_2$  and so on) as electron donors in anoxygenic photosynthesis, thus enhancing



A living, slightly squeezed *S. purpureum* cell (length about 40  $\mu\text{m}$ ) with symbionts, which in this case contain relatively few refractile storage vacuoles.

the efficiency of fermentative host metabolism by maintaining a low intracellular  $p_{\text{H}_2}$ . Excretion of photosynthates may also benefit the host. In the dark, photosynthetic nonsulphur bacteria are capable of oxidative phosphorylation using volatile fatty acids and hydrogen as substrates. Thus the symbiont has secondarily transformed the ciliate into an aerobic organism in the dark. The fact that, even in the dark, the ciliates tolerate only very low  $\text{O}_2$  tensions suggests that the symbionts protect the host against  $\text{O}_2$  toxicity by their respiratory activity.

The symbiotic relationship represents an analogy to the origin of mitochondria from photosynthetic bacteria. It shows how such association can be mutualistic under anaerobic conditions and how it can gradually convert an anaerobic host to aerobic life.

**Tom Fenchel**

**Catherine Bernard**

Marine Biological Laboratory,  
University of Copenhagen,  
DK-3000 Helsingør, Denmark

1. Young, D. et al. *Proc. natn. Acad. Sci. U.S.A.* **82**, 4443–4447 (1985).

2. Woese, C. R. *J. molec. Evol.* **10**, 93–96 (1977).

3. Fenchel, T. & Bernard, C. *FEMS Microbiol. Lett.* (in the press).

4. Fenchel, T. & Finlay, B. J. *Eur. J. Protist.* **26**, 210–215 (1991).

## Giant African lakes revisited

**SIR** — The idea that Southern Africa in the Upper Permian was a terrain of giant lakes, with herds of dicynodonts grazing by the lake shores (K. Yemane *Nature* **361**, 51–53; 1993) demonstrates that *plus ça change*... The geology of the Karoo was first studied in detail by A. G. Bain, a road engineer in Ft Beaufort, South Africa, 150 years ago.

Bain discovered the dicynodonts and correctly inferred their affinity to mammals. He compared them with modern hippos, and used stratigraphic, palaeobotanical and facies evidence to show that they lived by great lakes. In 1864 he visited London, where he was entertained by Owen (who had provided the formal descriptions of Bain's fossils) and Murchison (who supported Bain's lake theory, with discussion of Lake Victoria, which they thought was a relict feature). Murchison introduced Bain to Captain Speke, and Bain later also met the Prince of Wales. Bain's views on palaeogeography are summarized in an article from 1857: "the Dicynodon with its numerous and wonderful congeners are among the oldest of reptiles, and when alive were the undisputed occupants of the Great South African Lakes" (*E. Prov. Mon. Mag.*, 1857).

**Euan G. Nisbet**

Department of Geology,  
Royal Holloway College,  
University of London,  
Egham TW20 0EX, UK

## Hidden message

**SIR** — Using a sophisticated computer program to compare the entire protein-sequence database with the entire *Oxford Unabridged English Dictionary*, the longest words that Gonnet and Benner<sup>1</sup> could find contained only nine letters apiece. Two ten-letter words were found by extending the comparison to ten other European languages and Esperanto<sup>2</sup>.

Using my own SIDESWIPE algorithm (scanning I do, even searching when I've plenty else (that I should be doing)), I found an interesting seven-letter word in the sequence for the human TFIID TATA-box binding protein that reveals a previously unknown activity for the protein — ATPASE<sup>3</sup>.

**Stephen C. Harvey**

Department of Biochemistry,  
University of Alabama at Birmingham,  
Birmingham, Alabama 35294, USA

1. Gonnet, G. H. & Benner, S. A. *Nature* **361**, 121 (1993).

2. Jones, D. *Nature* **361**, 694 (1993).

3. Nikolov, D. B. et al. *Nature* **360**, 40–46 (1992).

### Scientific Correspondence

Scientific Correspondence is a relatively informal section of *Nature* in which matters of general scientific interest, not necessarily those arising from papers appearing in *Nature*, are published. Because there is space to print only a small proportion of the letters received, priority is usually given according to general interest and topicality, to contributions of fewer than 500 words, and to contributions using simple language.



# In peace and in war

Robert D. Holt

**Natural Enemies: The Population Biology of Predators, Parasites and Diseases.** Edited by M. Crawley. Blackwell Scientific: 1992. Pp. 576. £32.50, \$69.95 (pbk).

IN a 1946 review of predation and vertebrate populations, the wildlife biologist Paul Errington mused, "Whatever else may be said of predation, it does draw attention". In the near-half-century since, there has been an enormous amount of research on interactions between natural enemies (defined to include parasites, pathogens and herbivores, as well as predators) and their victims. It is easy to see why, for the *pas de deux* between enemies and victims permeates all levels of biological organization, from energy flow through ecosystems to the moulding of individual behaviour by natural selection, and is moreover central to many applied ecological problems.

Michael Crawley, the editor of this fine, timely volume, has assembled a roster of distinguished authors to examine our current understanding of the ecology of natural enemies. A core set of questions ties their contributions together. How do the behaviour, life histories and guild structure of natural enemies reflect their dependence on victims for sustenance? How important are natural enemies in limiting and regulating prey populations and in moulding prey phenotypes over evolutionary time? How often do such interactions foster dynamic stability, as opposed to cycles, chaos or local extinction? How are interactions between enemies and victims influenced by the broader web of interactions in the community?

Space precludes my dwelling on the virtues of each chapter, which are all well written and present fresh insights even when reviewing familiar material. A quartet of conceptual chapters sets the stage. Crawley lays out the basic elements of natural enemy-victim population dynamics, with enough background to make the volume accessible to students. Three chapters crisply outline the main tools used in evolutionary analyses of enemy-victim relationships: optimality models, comparisons among taxa and coevolutionary models. The book's core is a rich smorgasbord on the population dynamics and evolutionary ecology of a wide range of natural enemies. There are four chapters that deal with vertebrate predators, one with marine invertebrates, three with arthropod predators and parasitoids, two with parasites, and one with the three-way interaction between bloodsucking arthropods, vertebrate hosts and their

shared parasites.

There is much fascinating natural history here, replete with cogent reexaminations of many standard topics. I was particularly taken by S. Northridge and J. Beddington's reinterpretation of large size and blubber in marine mammals, the implications of these traits for whale species' coexistence, and how these ideas echo I. Hanski's elucidation of competition between shrews. There is also a chastening recognition of limits to our understanding. For instance, M. Hassell and C. Godfray succinctly review the literature on insect host-parasitoid dynamics and ruefully conclude that the "role of parasitoids in the regulation of natural insect populations remains uncertain".

Despite the great deal now known about natural enemy-victim interactions, my feeling is that we are still far from a general theory that would let us gauge accurately the ecological and evolutionary role of natural enemies in the play of life. This book reveals substantial agreement among ecologists as to the elements needed to address the above organizing questions, but in no empirical case are all the requisite components fully

understood. Yet some pointers toward a general theory do emerge, particularly in the concluding synthetic chapters, which emphasize the need for a better understanding of the mechanistic behavioural underpinnings of predator attack and prey defence, as well as the community context of these interactions. Two chapters underline the importance for human affairs of the material in the volume. The editor tops it all off with an overview of tentative generalizations and suggestions for future research.

I am struck by how certain themes pervading this volume reflect broader currents in ecology as a whole: the significance of scale and heterogeneity; the modest aim of developing contingent theory, tailored to particular systems, rather than a quixotic search for universal ecological principles; and the importance of broadly comparative, evolutionary perspectives in analyses of ecological systems. The book should be read, savoured and digested by all professional ecologists and applied biologists interested in its principal ecological theme, and particularly by conservationists concerned with the preservation of predators or their victims. For, as Finley Peter Dunne once said, "Life'd not be worth livin' if we didn't keep our inimies" (*Mr Dooley in Peace and in War*, 1898). □

Robert D. Holt is at the Museum of Natural History, University of Kansas, Lawrence, Kansas 66045-2454, USA.



**Web of intrigue** — an adult female garden spider (*Argiope appensa*) eating a honey-bee while a smaller male spider waits nearby. The picture is taken from *Hawaiian Insects and Their Kin* by F. G. Howarth and W. P. Mull, which contains over 200 colour photographs of such Hawaiian peculiarities as predatory happyface spiders and 'slinging' tree snails. University of Hawaii Press, \$19.95.

# The laughing physicist

Roland Pease

**The God Particle.** By Leon Lederman with Dick Teresi. *Houghton Mifflin*: 1993. Pp. 352. \$24.95. To be published in the UK by Bantam in June.

AS Americans brace themselves to spend several billion dollars on the Superconducting Super Collider (SSC), high-energy physicists have found themselves compelled to justify the expense. Steven Weinberg's recent *Dreams of a Final Theory* is a product of this new obsession (reviewed in *Nature* 361, 697; 1993).

Leon Lederman, one of the leading advocates of the SSC, takes on the same remit in his new book, *The God Particle*. The difference in the two authors' approaches can be seen in the subtitles of their books: *The Search for the Fundamental Laws of Nature* for Weinberg, with a series of low-key essays defending the high-energy physicist's view of the world; and *If the Universe is the Answer, What is the Question?* for Lederman.

This marks out Lederman as a joker before the cover is even opened. Indeed, he has quite a reputation for his gags. At first, the endless one-liners grate (the inevitable rabbi joke in Chapter 6 confirms a distinct resemblance to Woody Allen), but with acclimatization and a little restraint on the author's part, they soon become part of the easy-going style. Nonetheless, Weinberg is the clearer writer, which is a pity as Lederman in many ways has a more interesting story to tell.

This is because whereas Weinberg (along with Hawking, Davies, Close and other professional popularizers) is a theoretician, Lederman is an experimentalist. Although Weinberg's calculations may have shown the way to the SSC, Lederman's experiments have been among the ground-breakers that made it necessary. Experiments make the world tangible in a way theory cannot. Part of the joy of *The God Particle* is the way it describes the exhilaration of taking on the world experimentally to make it reveal its secrets (more than once Lederman likens the experience to sexual arousal).

Lederman's experimental credentials are powerful. A student working on Columbia University's Nevis Cyclotron when accelerator physics was establishing itself, he eventually became the director of the world's most powerful collider, the Tevatron at Fermilab, as well as a prime mover behind the SSC, now under construction in Texas. On the way, he picked up a Nobel prize for

physics for discovering that the electron- and muon-neutrinos are distinct particles, and could have picked up two more (don't take my word for it, read the book and take his) for work on parity violation and for discovering the long-lived kaon (which later revealed charge-parity, CP, violation).

Part history, part autobiography, part polemic, with sideswipes at mystics on the way, the book is a strange hotch-potch held together more by Lederman's outsized personality than any logic. He is hampered in part by the need to tell his story — the quest for the ultimate 'atom' — in a different way from his theorist



**Lederman: pipe dreams of a final theory?**

predecessors ("I feel like Zsa Zsa Gabor's seventh husband", he confesses early on, "I know what to do, but how do you make it interesting?").

He starts with Democritus — the 'Laughing Philosopher', appropriately enough — and the invention of the notion of the a-tom, the indivisible particle, outlined in the form of an imaginary nocturnal visitation and conversation with the great man. The circuitous route towards the atom takes in many of the usual characters, Galileo ("the Carl Sagan of the 1600s"), Newton, Dalton, Einstein and so on, but again with the experimentalist's peculiar insight.

One character you won't find in the other versions of this story is Roger Bosovich, an eighteenth-century Jesuit from Dubrovnik. His particular insight was to recognize that an atom must not be merely indivisible, but also dimensionless, reaching out into space by forces radiating from it. Time and again in this history, Lederman returns to his terrible twins, Democritus and Bosovich, and their criteria, to judge science's progress towards the God Particle.

And what is the God Particle of the title? At a time when obscurantists are finding happy offence in phrases such as Einstein's "mind of God" and George Smoot's "face of God", isn't Lederman

being rather provocative? He neatly justifies the term with the Old Testament story of the Tower of Babel. The God Particle is the Higgs boson, the principal quarry of the SSC, and just as God, with a kind of twist, turned one language into many, confounding our attempts to see heaven, so She (PC — political correctness — is not violated in this text) sent the Higgs boson to distort the underlying simplicity of the real world. Put another way: "before the Higgs, symmetry and boredom; after the Higgs, complexity and excitement". Like a patent medicine, the Higgs is advertised to cure all ills. Why are there muons as well as electrons? Send for the Higgs. Why do  $Z^0$ s have mass, but photons don't? Send for the Higgs. Why is there matter but not antimatter? Why did the Universe start with a moment of inflation? Send for the Higgs.

Sadly, Lederman then fails to give a convincing description of how the Higgs does all its herculean tasks: a trio of common examples of symmetry breaking (such as the magnetization of iron) and that's it. No metaphors for the mechanism, nor any description of how it will be recognized at the SSC.

And if the SSC does find this magic key, will that be the end of the story, the final theory? Apparently not. In the last chapter, Lederman imagines another conversation with Democritus, this time in the year 2020, and the location the next-generation accelerator where they are probing the intricacies of supersymmetry and grand unified theories. The quest continues. And has he justified the cost? Well even if he has failed to convey the excitement of the quest to all readers, after a decade's porkbarrelling, Lederman has an argument for the truly hardnosed: the search for the atom has already given us quantum theory, and that is responsible, he says, for an estimated 25 per cent of the gross national products of the industrialized nations. Is that God or mammon speaking? □

Roland Pease is assistant News and Views editor of *Nature*.

## New in paperback

■ *The Shadows of Creation: Dark Matter and the Structure of the Universe* by Michael Riordan and David Schramm. Oxford University Press, £9.99. For a review, see *Nature* 352, 677 (1991).

■ *The Mind of God: The Scientific Basis for a Rational World* by Paul Davies. Touchstone (Simon and Schuster)/Penguin, \$12/£6.99. For a review, see *Nature* 356, 632 (1992).

■ *In Search of the Edge of Time* by John Gribbin, an account of the science of time, time travel and time machines. Black Swan (Transworld), £5.99.



# Mind and meaning

John Fox

**Human and Machine Thinking.** By Philip N. Johnson-Laird. Lawrence Erlbaum: 1993. Pp. 189. £19.95, \$37.95.

"WHAT is thinking?" asks Johnson-Laird at the start of his engaging discussion of the similarities — or rather the dissimilarities — between current theories of machine reasoning and our understanding of the process of human thought. Johnson-Laird is well known for work that combines elegant experimental studies of human reasoning with careful consideration of ideas from mathematical logic. Here he argues that human reasoning differs markedly from that prescribed by classical deduction. This, his third book on the general topic (see *Mental Models*, Harvard University Press, 1983, and *Deduction* with R. M. J. Byrne, Lawrence Erlbaum, 1991), contains new material about non-deductive forms of reasoning and developments in artificial intelligence (AI).

Over the past 25 years or so, mathematical logic has increasingly provided the formal foundations for research into many topics in AI. Most major conferences on the subject have sessions on logical theory and technology, and there are many smaller meetings on specialized logics for problem-solving, planning and scheduling, natural language understanding and so on. The starting points for these developments were the classical propositional and predicate logics, but they have been augmented by countless new systems for reasoning about space and time, uncertainty and belief, actions and events, and many other aspects of what is often referred to as "common-sense reasoning".

In Chapter 1, the author summarizes much of his published work on human deduction, and provides some background on logic, together with consideration of less classical formal techniques recently developed in AI. The chapter provides a good restatement of his "mental models" theory of human reasoning which, he argues persuasively, accounts for many findings from laboratory studies of human thinking. The central claim is that logicians, computer scientists and most psychologists have concentrated on formal (syntactic) aspects of reasoning and neglected the true basis of human thought, which is grounded in an understanding of the meaning (semantics) of the material being dealt with and not just its form.

Deduction is often viewed as reasoning from the general to the specific. In

Chapter 2, Johnson-Laird discusses such human abilities as deducing general rules or laws from specific observations. Although there has been a vast body of work on 'plausible reasoning' in AI, a general inductive capability has not yet been achieved in machine reasoning. There has been progress on simple induction-like techniques but Johnson-Laird is unconvinced by their adequacy to account for interesting features of induction, rightly in my view. As with deduction, he places the blame for the absence of a good theory of inductive reasoning squarely on the lack of insight into how we actually understand the world. He presents a preliminary account of how model-based reasoning may relate to good methods of induction and discusses the causes of human error.

It is in the final chapter that Johnson-Laird deals with the most mysterious feature of human reasoning — creative thought. Building on a lengthy discussion of musical composition, and a somewhat more superficial discussion of scientific creativity, he arrives at the conclusion that there are several candidate mechanisms for creative thinking, and hints that the mental models approach will shed light on them, although the view is not well developed. His ideas are undoubtedly interesting, but this chapter is the least persuasive of the three. (One reason for doubting the adequacy of the heavily cognitive approach is that it neglects the sociological factors that seem so important in most kinds of creativity and innovation.)

Johnson-Laird's central thesis is that we need to underpin accounts of human and machine thinking with 'semantic' theories. My biggest worry is that he gives us little idea about what such theories might look like. Too much of his argument dwells on the weaknesses of classical logic, while ignoring recent ideas about common-sense reasoning and related subjects. Little of this AI work, on logics of time, space, belief, action and so on, has explicitly psychological motivations, but it should be considered for what it may tell us about our intuitive understanding of the world.

Nevertheless, the book is stimulating, instructive and enjoyable. The technical

discussions are well informed, and include clear presentations of important subjects for which the specialist literature is often inaccessible to the non-specialist. The volume deserves to be widely read. □

*John Fox is in the Advanced Computation Laboratory, Imperial Cancer Research Fund, PO Box 123, Lincoln's Inn Fields, London WC2A 3PX, UK.*

# In at the deep end

Damian Lawler

**River Mechanics.** By M. S. Yalin. Pergamon: 1992. Pp. 219. £35, \$70.

THE fluvial sciences are heavily based on the work of hydraulic engineers stretching back three centuries. M. S. Yalin, drawing on his many years of prolific research on problems of sediment transport and fluvial hydraulic engineering, has produced an extremely detailed, authoritative, research-level volume on selected aspects of the behaviour of alluvial channels.

The book is tightly focused, and moves logically from a consideration of flow processes (especially the recent exciting discoveries of bursting processes), through sediment transport theory, bedforms and regime channels, to meandering and braided river development. The approach has a strong theoretical and experimental base, but is also cross-disciplinary: the author successfully integrates many useful ideas and formidable data compilations from the increasingly linked fields of hydraulic engineering, geology and fluvial geomorphology.

The book is up to date and has a truly international perspective: there is a particularly successful synthesis of vast quantities of literature from North America, Europe, Japan and the former Soviet Union. Yalin is also delightfully thought-provoking (some may say controversial), especially in his discussion of scale effects in fluvial systems, and his stress on the importance of horizontal turbulent bursting phenomena in meander development, playing down the roles of secondary circulation and alternate-bar generation.

This is a condensed, selective book, however, and its tone makes few concessions to the nonspecialist reader. Also, Yalin's declared aim "to improve the understanding and formulation of the fluvial processes" suggests a breadth that is not there. For simplicity, he deals only with alluvial channels in cohesionless

■ Also just published on the subject of AI is *The Simulation of Human Intelligence* edited by Donald Broadbent. The volume is based on the 1991 Wolfson College Lectures and contains eight essays aimed at a general, nonspecialist audience by Roger Penrose, Allen Newell, Richard Young and Thad Polk, Dana H. Ballard, Edmund T. Rolls, Michael Brady, Gerald Gazdar, Margaret Boden and Donald Broadbent. Blackwell, £40 (hbk), £12.99 (pbk).

materials — principally large sandbed rivers: he neatly sidesteps, for example, the topical issues of geotechnical stability analyses, bank erosion mechanisms, riverine biological processes and solute and contaminant transport dynamics. The approach, too, can be narrow: the lack of a basin-wide, 'downstream' perspective is not unusual in hydraulic engineering analyses, but it is disappointing to find, for instance, the suspended sediment discussion couched simply in fluid energy terms, with no examination of the concepts of sediment supply, exhaustion and hysteretic behaviour, which we now know are fundamental to an understanding of the process dynamics. Similarly,

to base so much of the book on laboratory evidence and yet to omit discussion of the reliability of such hardware modelling is surprising. Also, some readers may not find the material very accessible because of a scanty index, a confusingly structured notation list and the curious absence of page numbers for references.

Nevertheless, this is an interesting book that clearly demonstrates the value of integrating fluvial research from the Earth and engineering sciences. □

*Damian Lawler is in the School of Geography, University of Birmingham, Birmingham B15 2TT, UK.*

## The jewel in the crown

*J. R. Tata*

**In Pursuit of Excellence: A History of the Indian Institute of Science.** By B. V. Subbarayappa. *Tata McGraw-Hill New Delhi, India: 1992. Pp. 426. \$11.50.*

THERE are few leaders in science and technology in India whose careers have not been moulded by the Indian Institute of Science. B. V. Subbarayappa has succeeded in producing a scholarly but readable account of the founding of this institution, its difficult early days and its blossoming. He also provides fascinating insights into colonial bureaucracy, personality clashes, intrigues and even some scandal. And in demonstrating what is possible in developing countries he has, perhaps unintentionally, revealed how not to start up a similar venture.

Three events in the nineteenth century were particularly important for higher education, science and technology in India: the acceptance of Lord Macaulay's recommendation in 1835 for the use of English in education; the establishment in 1857 of three universities at Calcutta, Bombay and Madras, modelled on the University of London; and the foresight and determination in the late nineteenth century of the Indian industrialist and philanthropist Jamsetji Nusserwanji Tata.

J. N. Tata was convinced that the only way to achieve the rapid industrialization of India was to provide a level of education and research in science and engineering that the newly founded universities, then mostly concentrating on the humanities, were incapable of. During his travels in Europe, the United States and Japan, he was impressed by post-graduate institutions of basic science and engineering and decided in 1896 to found with his own money a technological 'university' modelled on Johns Hopkins University in Baltimore, Maryland.

What then follows is a sad tale of protracted negotiations, incompetence and bickering among various authorities about the site and scope of the institute. After a two-month fact-finding mission in 1900, the southern city of Bangalore was eventually chosen as the site. Another four years were to pass before the organizational details were agreed on, by which time Tata was seriously ill. He did not live to see the birth of the Indian Institute of Science in July 1911.

The Scottish chemist Sir William Ramsay succeeded in persuading his colleague Morris Travers, then professor of chemistry at the University of Bristol, to accept the job as the institute's first director at a handsome annual salary of £1,800 and a life pension of £500. Travers had little taste for administration, nor was he selflessly interested in developing Indian science. Subbarayappa paints a bleak and often humorous picture of Travers's never-ending conflict with the council of the institute, the founder's family, builders and his own staff and students. Eventually, following accusations of embezzlement of institute funds, Travers was forced to resign as director, and returned to England in mid-1914.

A period of relative calm ensued, allowing the institute to consolidate its position, increase the number of students and start new departments, including India's first department of biochemistry; the institute's reputation continued to grow and it began to be taken seriously abroad. On M. O. Forster's retirement as director in 1932, it was decided that his replacement should be an Indian scientist.

The position went to C. V. Raman, the first Indian scientist to be awarded the Nobel prize, who was also appointed head of the department of physics. The department soon became internationally

renowned, but at the expense of the rest of the institute. Like Travers, Raman was a disaster as director. He was cantankerous, eccentric, quarrelsome and not interested in administration. The council, after the usual appointment of investigative committees, forced him to resign, despite Lord Rutherford's pleas to the then viceroy, Lord Linlithgow, to intervene.

The rapid expansion of the institute following Indian independence in 1947 was aided by the appointment of directors who were able administrators and by financial support from central agencies such as the Council of Scientific and Industrial Research and University Grants Committee. The institute was now in a position to recruit people of high calibre to head new departments, mostly engineering-based at first, resulting in the spawning of high-tech industries around Bangalore and, later, the forging of close links with the Indian Space Research Organisation. Expansion of life sciences came later still, attracting biotechnology-based institutions to sites near the institute (for example, the Astra Research Centre India). In the nine years since its platinum jubilee in 1984, the institute's annual budget, under the directorship of C. N. R. Rao, has increased more than fivefold, and the number of graduate students has grown to 1,500, with 100–125 PhDs being awarded each year. It could be argued that the institute is now spreading itself too thinly, but perhaps a more urgent concern is to lower the average age of its 475 academic staff.

Many of the 600 publicly funded scientific institutions in India are showpieces for the principal government agencies, with more glamorous buildings and being better funded and equipped than the Indian Institute of Science. But none enjoys the institute's prestige and reputation, nationally or internationally. One has only to look at the visitor's book to see what an attraction the establishment is. □

*J. R. Tata is at the National Institute for Medical Research, Mill Hill, London NW7 1AA, UK.*

### New editions

■ *An Introduction to Behavioural Ecology* by J. R. Krebs and N. B. Davies (3rd edn). Covers important new topics such as the latest techniques of comparative analysis and DNA fingerprinting. Blackwell Scientific, £17.95 (pbk).

■ *Biogeography: An Ecological and Evolutionary Approach* by C. B. Cox and P. D. Moore (5th edn). Now addresses the effects of current human activities and the prospect for the future. Blackwell Scientific, £16.50 (pbk).



# Optimality, mutation and the evolution of ageing

L. Partridge & N. H. Barton

**Evolutionary explanations of ageing fall into two classes. Organisms might have evolved the optimal life history, in which survival and fertility late in life are sacrificed for the sake of early reproduction and survival. Alternatively, the life history might be depressed below this optimal compromise by deleterious mutations: because selection against late-acting mutations is weaker, these will impose a greater load on late life. Evidence for the importance of both is emerging, and unravelling their relative importance presents experimentalists with a major challenge.**

DESPITE all human efforts to the contrary, death is one of life's only two certainties. It is often preceded by intimations of mortality, in the form of a decline in fertility and most aspects of biological performance, characteristic of 'senescence' or 'ageing'<sup>1-4</sup>. Ageing presents not only a medical problem, but also an evolutionary paradox: if organisms can function well in youth, why can they not continue to do so in old age? Recent theoretical work has clarified two possible causes of ageing. It could evolve as the necessary cost of processes beneficial to youth, or could instead be purely maladaptive, and evolve because of the pressure of deleterious mutations on populations<sup>2,4</sup>. Evidence for the importance of both causes of ageing is emerging, and unravelling their relative importance presents experimentalists with a major challenge. The view that mutation contributes to the senescence of individuals parallels recent emphasis on its role as an evolutionary force capable both of maintaining sexual reproduction and of producing decline and extinction of evolutionary lineages<sup>5,6</sup>.

We confine the term 'ageing' or 'senescence' solely to the drop in survival probability and/or fertility later in the life of individuals, to distinguish it from other changes during the life history, such as early development or the onset of reproduction. Ageing reduces the contribution that older individuals make to future generations. It is often seen in nature<sup>7-9</sup>, but becomes much more obvious when natural hazards are removed in captive and human populations<sup>1,10</sup>. The rate of ageing is highly variable. In semelparous forms, which breed only once, there is often a sudden loss of function after reproduction; adult mayflies live for less than a week, and male Pacific salmon show catastrophic senescence after their single breeding attempt. In iteroparous organisms, which can breed repeatedly, the decline is usually more gradual, as in many beetles and in fish such as haddock and rockfish.

Mechanistic accounts of ageing invoke various forms of damage, to DNA, cells, tissues and organs<sup>3,11-14</sup>. From this perspective, the rate of ageing could be determined solely by level of exposure to damaging influences: the process would be inevitable, and in need of no evolutionary explanation. But even under optimal conditions in captivity, we see very different potential lifespans in organisms that do not seem to differ in their risk of damage, showing that the degree of ageing has evolved. For instance, most birds considerably outlive mammals of comparable size. Among more closely related organisms, bats have very much longer maximum lifespans than do similar-sized rodents, and the flightless emu and ostrich are notably short-lived among birds<sup>1,10</sup>. The thick-shelled bivalves have greater lifespans than do other molluscs, and turtles and tortoises outlive other reptiles<sup>10</sup>. The proximate cause of ageing may well be various kinds of damage; however, these comparisons show that organisms vary in the extent to which they avoid or combat it<sup>10,15</sup>. It is this variation that any evolutionary theory of ageing must explain.

## Evolutionary theories of ageing

Because ageing reduces the genetic contribution of individuals to future generations, it is opposed by natural selection. But as Medawar first pointed out<sup>16,17</sup>, the natural selection that maintains survival and fertility becomes weaker through the life history. Even without ageing, organisms are at risk of death and impaired fertility from disease, predation and accidents. If the genes affecting survival and fertility are to some extent age-specific in their effects, then those that influence later life will be subject to weaker selection because, by the time they take effect, more of the original carriers will already have died or become infertile for other reasons<sup>10,16,17</sup>. (Note that what matters is when a gene affects survival and fertility, not when it is expressed; for example, a gene which is expressed early in the development of the heart might increase the chance of heart disease much later in life.) The sensitivities of fitness to changes in survival and fertility at different ages are explained in Box 1, and illustrated by Fig. 1, which is based on data from lines of *Drosophila* selected for early or late reproduction<sup>18,19</sup>.

We shall contrast two kinds of explanation for the evolution of ageing. First, constraints on the combinations of survival and fertility that the organism can achieve at each age may mean that a single optimal genotype evolves, which shows senescence because fitness is maximized by increasing early performance at the expense of late (the optimality explanation). Second, senescence might evolve because of a greater mutation load on the later, and less strongly selected, part of the life history (the mutational explanation).

The arguments over the relative importance of optimality versus mutational explanations are relevant to a wide range of evolutionary questions<sup>6</sup>, and the problem of senescence is one instance of a general question: to what extent does the degree of adaptation reflect the strength of selection? For example, does the precision of mimetic patterns in butterflies reflect the degree of protection they confer<sup>20</sup>. Can very weak selection maintain adaptations? Perhaps the best example of an adaptation maintained by weak selection ( $s \approx 10^{-5}$ ) is the bias towards use of the most efficiently translated triplet codon, seen for some loci in *Escherichia coli*, yeast and *Drosophila*<sup>21</sup>. Yet, we are still ignorant of the general magnitude of the selection that produces and maintains adaptations.

## Optimality theories of ageing

Optimality theory<sup>22,23</sup> has been successful in explaining diversity in life-history traits such as age of first breeding, number of breeding attempts and number of offspring produced at each<sup>2,24-27</sup>. Optimality models do not contain specific genetic parameters, and instead specify the life history that gives the greatest fitness, within the intrinsic constraints of physiology and the extrinsic constraints imposed by the environment. Optimality arguments are consistent at least with simple genetic

models<sup>28-30</sup>: the optimal life history is an evolutionarily stable strategy, which cannot be displaced by any feasible genotype.

Ageing could evolve as part of an optimal life history. On this view, senescence arises from the deleterious side-effects late in life of processes that are favourable early on (see Fig. 2). An antagonism between early performance and resistance to ageing can occur if prolonged growth during the pre-adult period elevates juvenile mortality but yields a longer-lived adult<sup>18,19,31</sup>. Similarly, early reproduction may impair survival or future fertility, by consuming resources, causing somatic damage or exposing the organism to environmental injury: there is a 'cost of reproduction'<sup>16,17,32-38</sup>. The 'disposable soma' theory is an optimality account of ageing, in which allocation of resources to reproduction jeopardizes somatic repair mechanisms and hence longevity<sup>15,39</sup>. Selection intensity on early, beneficial effects will be stronger than on later, deleterious ones and so the optimal life history will include ageing.

Of course, as for any other trait, it may be impossible for any one homozygous genotype to achieve the optimum. In diploids, it may instead be best approximated by a heterozygote, if each homozygote is better for a different life history trait, and the alleles are dominant for their advantageous trait (sometimes

known as 'antagonistic pleiotropy'<sup>40-42</sup>). If such balancing selection maintains a substantial fraction of life-history variation, it will generate negative genetic correlations among life-history traits. However, such polymorphism does not account for senescence itself, which depends on the mean life history, determined by maximization of fitness subject to what is possible.

The optimality account of ageing has come to be known as the 'pleiotropy theory of senescence', because it is often developed in terms of genes with effects on more than one aspect of the phenotype, in this case on survival and fertility at different ages. But mutational theories of ageing can also involve pleiotropy (for example see Box 2). Moreover, under the optimality theory the outcome does not depend on patterns of pleiotropy, and could in theory be reached by fixation of alleles each with an effect on fitness at only one age. (On Fig. 2, this could be represented by evolution up to the optimum life history by a series of horizontal or vertical steps.) Such evolution would imply that the constraints on optimization occur only on the trade-off curve itself, and not inside it, and so is hardly plausible: our point is that the optimality argument is independent of genetic details such as pleiotropy. Provided that the population is not trapped at a local optimum, and provided there is at least

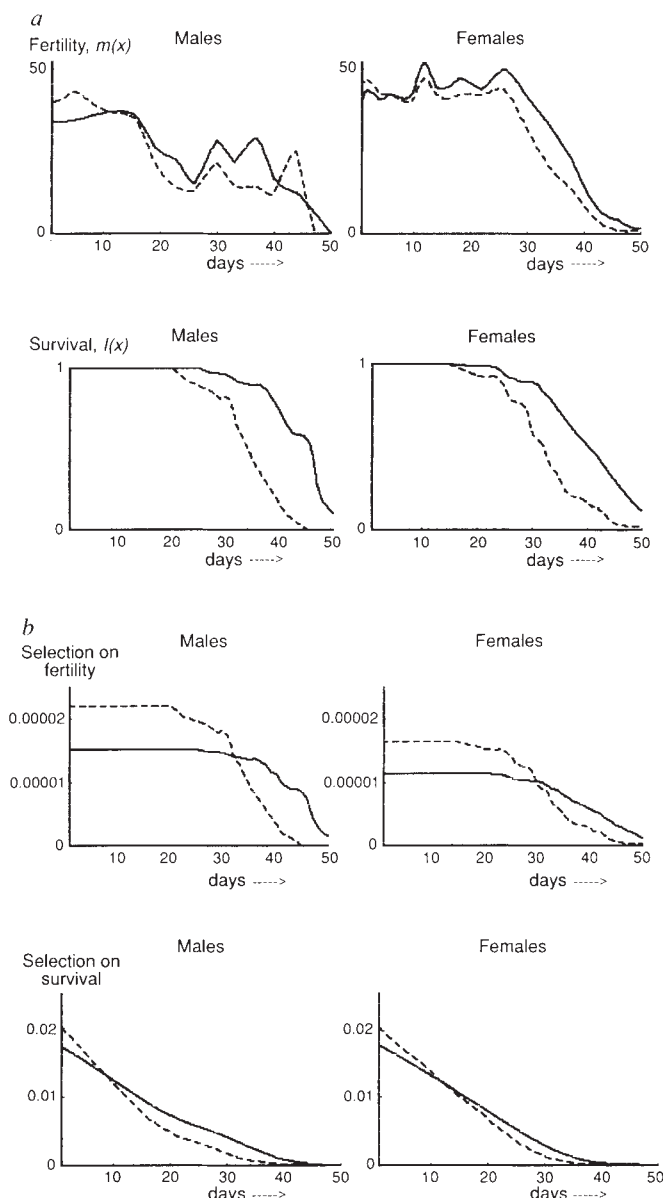


FIG. 1 An illustration of how selection on fertility and survival for the two sexes decreases with age. *a*, Survival,  $l(x)$ , and fertility,  $m(x)$ , in the two sexes, as a function of time since eclosion, for lines of *Drosophila melanogaster* which had been selected for early or late reproduction<sup>18,19</sup>. Female fertility is measured in eggs per day, and male fertility is measured as the percentage of progeny relative to a standard competitor. Dashed lines correspond to 'early' lines, solid lines to 'late' lines. Data are from crosses between replicate selected lines, to avoid inbreeding effects<sup>19</sup>. *b*, Sensitivity of fitness ( $r$ ) to survival and fertility in the two sexes; this is calculated from  $l(x)$ ,  $m(x)$  using equation (2) (modified to allow for separate sexes). In all four figures, selection is stronger in early life on the lines selected for early reproduction, (dashed lines) and stronger in late life for the 'late' lines. The difference is greater for selection on fertility (top row) than on survival (bottom row). The top pair of figures shows the sensitivity of fitness to an increase in fertility in the two sexes ( $\partial r / \partial m(x)$ ). We assume that larvae take 10 days to develop (ignoring differences between lines), and that larval viability is adjusted by density-dependence to maintain constant population size ( $r=0$ ). The units for females are rate of increase per day per extra egg; to translate these into selection coefficients per generation per extra egg, multiply by the generation time ( $T=25$  days for 'early' lines, 24 days for 'late'). Units for relative male mating success are as in *a*. The second pair of figures shows the sensitivity of fitness to a decrease in death rate, as a function of adult age ( $-\partial r / \partial \mu(x)$ ; equation (2b)). Units here are rate of increase per day per unit decrease in death rate. (Note that these figures are for a hypothetical population with the life history given by *a*, not for the life history as it was while the lines were selected.)



some genetic variation in the appropriate direction, the optimal life history will eventually be reached.

### Senescence as a consequence of mutation pressure

Mutation pressure could lead to ageing<sup>2,17,43</sup>. Because the intensity of selection on later-acting mutants declines with age, alleles with deleterious effects will reach a higher frequency in a mutation-selection balance the later the age at which they reduce fitness. In a small population, or with asexual reproduction, random sampling drift may combine with mutation to overwhelm selection, so that alleles with late, deleterious effects become common. Even in a large sexual population, senescence can be caused by the cumulative effects of many rare deleterious alleles held in mutation-selection balance.

With asexual reproduction of diploids, mutation reduces mean fitness by a factor  $\exp(-2U)$ , where  $U$  is the total number of deleterious mutations per haploid genome per generation, and mutations are assumed not to be completely recessive<sup>44,45</sup>.

#### BOX 1 Fitness in age-structured populations

THE dynamics of an age-structured population depend on the probability  $l(x)$  of surviving from birth to age  $x$ , and the expected number of offspring at that age,  $m(x)$ . The dynamics depend only on the product of survival and fertility, which we denote by  $k(x) = l(x)m(x)$ . A homogeneous population will eventually grow at a steady rate  $r$ , which is given by the Euler-Lotka equation,

$$1 = \int_0^{\infty} \exp(-rx)k(x) dx \quad (1)$$

In an asexual population, or a population of sexually reproducing haploids that vary only at a single locus, the outcome of natural selection depends simply on the long-term growth rates associated with each genotype, in the absence of density- or frequency-dependent interactions, each genotype will eventually grow exponentially at a rate that depends on its own life history, given by equation (1). With certain simplifying assumptions (that selection is weak, and mating is random with respect to age and genotype) the theory can be extended to diploids and to variation at multiple loci<sup>2</sup>.

The intensity of selection on different stages of the life history can be found from the effect on fitness,  $r$ , as defined above, of small changes in fertility or mortality at age  $x$  (refs 2, 32)

$$\frac{\partial r}{\partial m(x)} = \frac{\exp(-rx)l(x)}{T} \quad (2a)$$

$$\frac{\partial r}{\partial \mu(x)} = \frac{-\int_x^{\infty} \exp(-ry)l(y)m(y) dy}{T} \quad (2b)$$

where the death rate  $\mu(x)$  is defined by  $l(x) = \exp(-\int_0^x \mu(y) dy)$ , and  $T = \int_0^{\infty} \exp(-rx)k(x) dx$  is the generation time. These expressions give the effect on fitness of small changes confined to a specific age,  $x$ ; they are applied to *Drosophila* data in Fig. 1. Similar expressions can be found for the effect of changes in age of onset of alleles which have permanent effects once first expressed<sup>32</sup>.

These formulae involve the rate of population growth,  $r$ ; this is because the value of an offspring depends on the size of the population into which it is born. In an increasing population, offspring produced early will themselves start to reproduce, and so are more valuable than those produced late. But in the long term, density-dependent factors must regulate population size by decreasing survival and fertility such that  $r=0$ . Then (for a given total reproductive output  $\int_0^{\infty} l(x)m(x) dx$ ), the timing of offspring production is selectively neutral.

Sensitivities of fitness to changes in both fertility and mortality are related to the probability of reaching the age of gene expression, and are inversely proportional to the generation time. For a gene affecting survival, the proportion of total reproduction expected to occur after that age (which depends on the fertility schedule) is also important (hence the integral in equation (2b))<sup>32</sup>. For mortality, the intensity of selection in a stable population ( $r=0$ ) remains constant during the pre-reproductive period because the consequences of death for reproductive success remain constant; selection intensity starts to decline only at the time of first breeding. □

Surprisingly, minor mutations reduce fitness just as much as those with larger effects, because they rise to higher frequency in a mutation selection balance<sup>46</sup>. This is why the mutation load (that is, the reduction in mean fitness) depends on the total mutation rate,  $U$ , and not on the strength of selection. If the effects of mutations on fitness are multiplicative, similar relations hold in sexual populations: fitness is reduced by  $\exp(-U)$  if mutations are recessive (or almost so), and  $\exp(-2U)$  otherwise<sup>47</sup>. Suppose now that deleterious mutations are age-specific. Then, the reduction in fitness caused by a loss in performance at various ages reflects the mutation rates to genes affecting those ages. But because a given change in performance has less effect on fitness later in life, we expect deleterious mutations to cause a greater drop in fertility and survival probability later in life, assuming the same mutation rate for each age. This increased load on later life can be accentuated by a positive feedback: as late-life performance declines, selection to maintain it becomes weaker, accelerating the decline. If mutations that specifically damage survival or fertility of the old are common enough, mutation may overwhelm selection<sup>2,10,16,30</sup>, causing catastrophic senescence (Box 2 and Fig. 2). The same argument can explain why different functions often decline at a similar age: synchronous collapse does not imply a single mechanism of senescence (J. Maynard Smith, personal communication).

Other evolutionary theories of ageing have been proposed. Arguing from the example of Huntington's chorea, Medawar<sup>17</sup> suggested that there could be selection for genetic modifiers delaying the age of onset of the effects of deleterious alleles. The argument was not that mutation itself caused senescence, but rather that, given some mutation load, the adaptive evolution of modifiers would shift the effects of that load to later life. This

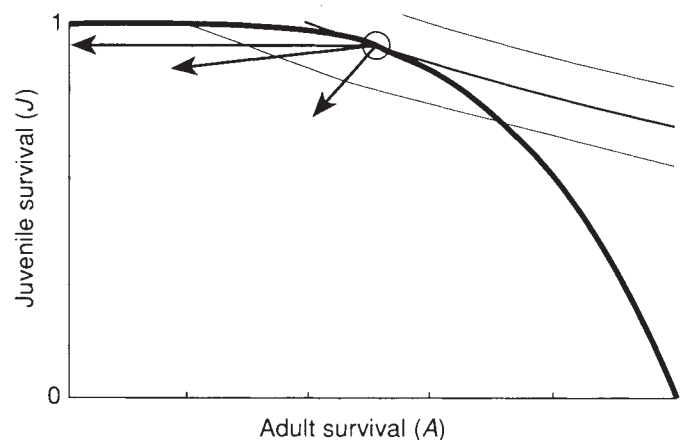


FIG. 2 A simple model of an organism with two age classes (see Box 2). Each female produces one daughter at each of two ages ( $m=1$ ); the chance of survival from birth to age 1 is  $J$  and from age 1 to age 2 is  $A$ . The heavy curve shows the maximum possible survival probabilities, and represents the trade-off between adult and juvenile survival ( $J+A \leq 1$ ). The light curve shows contours of fitness, spaced at intervals  $\Delta r=0.1$ . The optimal life history is where the furthest contour from the origin just touches the trade-off curve (open circle,  $J=0.935$ ,  $A=0.505$ ). The arrows show the reduction of survival below the optimal value caused by deleterious mutations at a rate  $U=0.1$ . The horizontal arrow shows the effect of mutations that reduce only adult survival ( $\delta J=0$ ;  $v=(\delta A/A)/(\delta J/J)=\infty$ ); when the net rate of mutation exceeds a threshold ( $U>0.067$ ), these accumulate indefinitely, and the life history becomes semelparous. The lower diagonal arrow shows the effects of mutations that reduce juvenile and adult survival equally ( $\delta A=\delta J$ ;  $v=1$ ). There is now no threshold, as the net mutation rate,  $U$ , increases, there is a steady decrease in survival, to 0.765, 0.414 when  $U=0.1$ . The intermediate arrow shows the effects of mutations with a slight pleiotropic effect ( $v=16$ ); these reduce survival to 0.866, 0.155 when  $U=0.1$ .



is closer to the optimality theory than to one based on straightforward mutation accumulation. As for Fisher's analogous explanation of the evolution of dominance<sup>48</sup>, the argument is implausible, because selection on such specific modifiers is extremely weak (of the order of the mutation rate per locus) and likely to be overwhelmed by their pleiotropic effects<sup>2,49</sup>.

Hamilton<sup>32</sup> argued that, because selection is stronger on earlier ages, alleles with advantageous effects early in life would be more likely to be fixed by selection than those with a later age of action, an argument that implies that populations are continually adapting, and so are not in equilibrium. The late part of the lifespan would then lag behind the current optimum more than would the early part, and senescence would be caused by the extra 'lag load'<sup>50</sup>. This idea can be seen as the converse of the mutational argument. For it to be plausible, continual adaptation must be necessary for the maintenance of the life history, just as the Red Queen had to keep running to stay in the same place<sup>51</sup>. It is then the failure to keep up with the requirements of changing conditions that leads to a substantial 'lag load' on late life. The idea is perhaps implausible because it does not explain why ageing is associated with pathological

changes, which would have been the ancestral state under Hamilton's theory.

## Empirical tests of evolutionary theories of ageing

We must distinguish between predictions that are common to all evolutionary theories, and those that can distinguish between optimality and mutation-accumulation. Some support for evolutionary theories comes from the comparative findings, which show that high maximum lifespans in optimal conditions (cited in the introduction) occur in creatures with a low risk of mortality in nature, and hence with strong selection maintaining late life. For example, birds have higher maximum lifespans than mammals and are less prone to death in the wild. The ability to fly and so escape predators may be important: among mammals, bats survive well both in nature and under optimal conditions, whereas among birds flightless species survive poorly in both situations<sup>1,7,9,10,52,53</sup>. Bivalves and tortoises have longer maximum lives than their relatives, perhaps because of the protection conferred by their thick shells<sup>1,10</sup>. High maximum lifespans are also found in organisms whose fertility increases with age, such as many fish and trees, as the evolutionary theory

### BOX 2 Optimality and mutation

HERE we illustrate optimality arguments, and also show how mutation pressure can lead to catastrophic senescence. Consider a simple model of an organism that reproduces at just two ages. The probability of survival from birth to age 1 is  $J$ , and from age 1 to age 2 is  $A$ . For simplicity, we assume that the expected fertility is the same at both ages ( $m(1)=m(2)=m$ ; hence,  $k(1)=mJ$ ,  $k(2)=mJA$ ). The growth rate is then given by the discrete version of equation (1),

$$1 = mJ e^{-r} + mJA e^{-2r} \quad (3a)$$

This has solution:

$$e^r = \frac{mJ}{2} \left( 1 + \sqrt{1 + \frac{4A}{mJ}} \right) \quad (3b)$$

First, consider the optimal life history, given some constraint on the rates of juvenile and adult survival that can be achieved. We suppose that it is impossible to sustain both high juvenile and adult survival, and choose the arbitrary constraint that  $J + A^4 \leq 1$ . Then there are progressively large decreases in adult survival for each increment in juvenile survival, and hence a convex trade-off curve (heavy line in Fig. 2). Other constraints give qualitatively similar conclusions, provided that the curve remains convex. For fertility  $m=1$ , the life history that maximizes fitness ( $r$ ) involves substantial senescence, in the sense that adult survival is much lower than juvenile,  $J=0.945$ ,  $A=0.505$  (open circle in Fig. 2).

Now, suppose that deleterious mutations reduce survival below this optimal value. Consider a class of mutations that reduce juvenile survival by a factor  $(1-\epsilon)$ , and adult survival by a factor  $(1-v\epsilon)$ , where  $\epsilon \ll 1$ ;  $v$  describes the degree of pleiotropy. These will reduce fitness by,

$$\begin{aligned} \delta r &= \epsilon \frac{\partial r}{\partial \mu(1)} + v\epsilon \frac{\partial r}{\partial \mu(2)} \\ &= \frac{-\epsilon(1+2(v+1)\gamma + \sqrt{1+4\gamma})}{\sqrt{1+4\gamma}(1+\sqrt{1+4\gamma})} \end{aligned} \quad (4)$$

where  $\gamma = A/(mJ)$ . ( $\delta r$  is negative because these mutations decrease fitness). Selection on deleterious alleles depends on the current life history through  $\gamma$ . When fertility and adult mortality are high ( $m, J \gg A$ ;  $\gamma \approx 0$ ),  $\delta r \approx -\epsilon$ , so that selection depends on the effects on juvenile survival. When fertility is low and juvenile mortality is high ( $m, J \ll A$ ;  $\gamma \gg 1$ ),  $\delta r \approx -(v+1)\epsilon/2$ , and selection depends on the average effect on juvenile and adult survival.

If we just consider a single locus, the equilibrium frequency is  $p = \mu/\delta r$ . For small  $\mu$ , homozygotes are rare, and the frequency of heterozygotes is approximately  $2p$ .  $J$  and  $A$  are therefore reduced by factors

$$\begin{aligned} (1-2\mu\epsilon/(-\delta r)) &\approx \exp(-2\mu\epsilon/(-\delta r)) \\ (1-2\mu\epsilon v/(-\delta r)) &\approx \exp(-2\mu\epsilon v/(-\delta r)) \end{aligned}$$

respectively. The next step is to combine these factors across all loci. For simplicity, we assume that the effects of different loci multiply, and we neglect linkage disequilibrium and genetic variation<sup>6</sup>, to derive approximations to the survival probabilities,

$$\begin{aligned} J &= \bar{J} \exp\left(\frac{2U}{\partial r/\partial \mu(1) + v \partial r/\partial \mu(2)}\right) \\ &= \bar{J} \exp\left(\frac{-2U\sqrt{1+4\gamma}(1+\sqrt{1+4\gamma})}{(1+2(v+1)\gamma + \sqrt{1+4\gamma})}\right) \end{aligned} \quad (4a)$$

$$\begin{aligned} A &= \bar{A} \exp\left(\frac{2vU}{\partial r/\partial \mu(1) + v \partial r/\partial \mu(2)}\right) \\ &= \bar{A} \exp\left(\frac{-2vU\sqrt{1+4\gamma}(1+\sqrt{1+4\gamma})}{(1+2(v+1)\gamma + \sqrt{1+4\gamma})}\right) \end{aligned} \quad (4b)$$

where  $\gamma = A/(mJ)$ , and  $\bar{J}$ ,  $\bar{A}$  give the optimal life history in the absence of mutation. These two equations can be solved numerically to find the equilibrium load,  $(J/\bar{J})$ ,  $(A/\bar{A})$ .

Two key points emerge. First, consider mutations that only affect adult survival ( $v=\infty$ ). As these increase in frequency, adult survival drops, and so selection against the mutations decreases. Above a critical threshold, mutation overwhelms selection, and the population becomes semelparous. Then, the only solution to equations (4) is at  $A=0$ ,  $J=\bar{J} \exp(-2U)$ . For the present example, the threshold is quite low,  $U=0.067$ . This quantifies the feedback effect postulated by Williams and Medawar<sup>10,20</sup>.

If, however, mutations affect both juvenile and adult survival equally, selection against their early effects keeps them at low frequency, and prevents the collapse of late survival. Surprisingly little pleiotropy is needed to reduce the degree of senescence substantially. For example, if mutations reduce adult survival 16 times more than they reduce juvenile survival (that is,  $v=16$ ), a net mutation rate  $U=0.10$  reduces juvenile and adult survival from their optimal values (0.945, 0.505) to (0.866, 0.155); if the mutations only affected adult survival, a mutation rate  $U>0.067$  would reduce adult survival to zero. Thus, the plausibility of mutation accumulation as an explanation of senescence depends on whether mutations with very age-specific effects are sufficiently frequent.  $\square$

\* Combining the effects of many loci is nontrivial. First, even though the effects of mutations on  $J$  and  $A$  are assumed multiplicative, their effects on fitness are not. Hence linkage disequilibrium may build up, and may substantially affect the equilibrium mutational load<sup>6</sup>. Second, mutation will generate variation in life history around the mean. A full treatment, following Charlesworth<sup>30</sup>, would be valuable.



of ageing would again predict. More data of this kind would be valuable.

It is very much harder to evaluate different evolutionary theories. Because we know that there is a substantial input of deleterious mutations to populations<sup>6,54</sup>, and that it is impossible for individuals to combine indefinitely high survival and fertility, both the optimality and the mutation-accumulation explanations of senescence must apply. The aim is to discover how far mutation has depressed survival and fertility below their optimal values. But determining the optimal life history of any organism is difficult because the trade-off curves that reflect constraints on survival and fertility at different ages have never been measured<sup>37</sup>. Furthermore, the total rate of production of deleterious mutations and their pattern of age-specificity are unknown<sup>6,55</sup>.

The amount of standing genetic variation for ageing within populations, and the number of genes involved, are not informative. Both theories are consistent with the absence of any genetic variance because deleterious mutations could become fixed by drift, and the optimal life history could in theory be achieved by fixation of a single optimal genotype by selection. By contrast, both theories are consistent with the presence of genetic variation for ageing, which is the norm for quantitative characters including other life-history traits<sup>56</sup>. At least in *Drosophila*, ageing is in general genetically variable and polygenic<sup>57-59</sup>. Genetic variation could result from a balance between mutation and selection, or from balancing selection alone; given the general difficulties of finding the causes and genetic basis of polygenic variation<sup>60</sup>, it is hard to take this line of evidence further.

The pattern of standing genetic variance has been exploited in investigations of the evolution of ageing. If senescence is primarily due to late-acting deleterious mutations, then (other things being equal) one would expect the additive genetic variance for survival and fertility to increase with age. One study<sup>61</sup> found no increase in additive genetic variance for female fertility with age in *Drosophila*. A problem is that mean values for fertility and survival probability decline with age, which may not necessarily be accounted for by a logarithmic transform. Also, some genotypes are lost from experiments through death, which will cause an inevitable decrease in variance if fecundity and survival are correlated.

Standing genetic variance has also been used to study the sign and magnitude of genetic correlations between ages, for survival probability and fertility. Results of chromosome extractions and breeding designs have been mixed, with some showing negative correlations between early and late fitness and others showing no significant associations<sup>62-65</sup>. These tests are not particularly sensitive; very large sample sizes are needed to generate reasonable confidence limits on genetic correlations<sup>66</sup> and biases can be introduced by deaths before late fitness is measured. In addition, some studies used inbred lines of *Drosophila*, which are liable to give artefactual positive correlations between fitness components, and are therefore unsuitable for testing evolutionary theories of ageing<sup>67</sup>.

Artificial selection has also been used to investigate the evolution of ageing. Restriction of breeding to older adults would be expected to reduce ageing because there is selection for higher longevity and high fertility at old ages in these 'old' lines. If a response to selection occurs, and if ageing in the original base population were attributable entirely to the presence of more or less age-specific deleterious mutations, then no immediate drop in survival or fertility would be expected to occur earlier in life. If trade-offs were important, then an immediate drop (a correlated response) would be predicted. This type of experiment in *Drosophila* has in general produced evidence for trade-offs, with lifespan and fertility late in life increasing in lines propagated from old adults, and either pre-adult survival or fertility of young adults showing a correlated decline<sup>18,19,68,69</sup>. Although the timescale of artificial selection experiments means that new mutations can contribute to the responses, in the above

experiments accumulation of partially recessive mutations could not account for the decline in early performance, because the drop persisted in crosses between 'old' lines<sup>19,58,59</sup>. In any case, the effect occurred too rapidly to be accounted for by mutation accumulation<sup>70,71</sup>. At least some of the increase in late life fitness in 'old' lines could have been caused by reduction in the frequency of predominantly late-acting mutations, and a study of the effects of reverse selection suggested that this was the case<sup>72</sup>, and that mutation accumulation had therefore also contributed to ageing in the base stock.

A fundamental difficulty with any approach based on standing genetic variance is that there need be no direct connection between the pattern of genetic covariances, and either the trade-off curve (which reflects the set of possible phenotypes), or the effects of spontaneous mutations. When variation in a set of quantitative traits is maintained by a balance between mutation and selection, the genetic covariances depend in a complex way on the covariances of effects of new mutations on the traits and on fitness, and on the pattern of selection on the traits<sup>30,73-76</sup>. Similar arguments apply when variation is a pleiotropic side-effect of balancing selection<sup>77</sup>. Artificial selection may circumvent some of these difficulties. Genetic variation within a base population can be magnified as variation between selected lines, giving greater statistical power. More important, selection over tens of generations can take traits well beyond their original values, and new mutations can contribute to the response, so that it may reflect the fundamental limits on the character, rather than just the initial covariances.

Artificial selection is not free of problems. It is possible inadvertently to apply direct selection to the characters whose correlated responses are of interest. For instance, in many experiments there may have been unintended selection pressure for rapid development and reproductive maturation in the control line flies with which the 'old' lines were compared, potentially leading to artefactual effects on early fertility<sup>19</sup>. In addition, all estimates of quantitative genetic variables by any method are sensitive to the environment in which they are measured; misleading gene-environment interactions can occur in an environment other than that in which the life history evolved<sup>78-80</sup>. This objection may apply to some extent to all published studies. Unless the base stock was kept for many years in the same conditions as those in which estimates were made, problems could have arisen.

An alternative to genetic approaches is to use environmental manipulations of the rate of ageing to study its mechanisms. For instance, in *Drosophila* females, reduction of egg production by X-ray irradiation, lowering of protein intake or denial of oviposition sites, has been shown to extend lifespan<sup>81-83</sup>. Some aspects of mating itself also shorten female lifespan<sup>84</sup>. These data suggest that both egg production and mating could accelerate ageing. Environmental manipulations have the considerable advantage that they can be used in wild populations, and they are also relatively quick. In addition, reversal experiments can be done, allowing the time course of the effects of the manipulation to be observed<sup>85</sup>. They can be especially powerful when used in conjunction with genetic manipulation<sup>86,87</sup>, and the two techniques are at their most convincing when they yield similar results.

In principle, one could find the contribution of deleterious mutations to ageing by turning off mutation, and allowing the population to recover to its optimal life history. Unfortunately, this is hardly feasible. One could instead increase mutation rates artificially: if doubling mutation rates doubled senescence, or if a slight increase in the mutation rate produced catastrophic senescence (Box 2), then this would suggest a substantial effect of mutation in the original population. This is a worthwhile approach, though one should bear in mind that different mutagens give qualitatively different kinds of mutation. Another promising way forward would be to use mutation-accumulation experiments<sup>54</sup> to make better estimates of the total mutation

load and its pattern of age-specificity. These experiments somewhat underestimate mutation rates because, even where new mutations are held heterozygous against balancer chromosomes, it is impossible to eliminate natural selection altogether if mutations are not completely recessive; to take the extreme example, dominant lethals can never be accumulated. But because only a small part of the mutation load is likely to be due to alleles with major effects<sup>6,54,71</sup>, this gives little error.

Long-inbred populations might be useful for assessing the effects of new mutations. In outbred populations, selection acts on new mutations mainly through their heterozygous effects. In contrast, under persistent inbreeding, the mutation load will mainly be due to homozygotes, and can be largely eliminated by outcrossing. This would allow direct measurement of how far senescence in the original population had been caused by accumulation of partially recessive mutations. Selfing populations of plants could be valuable material<sup>88,89</sup>, although the retention of unused yet presumably expensive floral features in some populations raised doubts as to whether they have been selfing for long enough to reach evolutionary equilibrium. An alternative might be to investigate the effects on life history of making a diploid from two independent strains of a habitually haploid organism<sup>90</sup>.

Additional evidence for the importance of mutation accumulation has come from experiments where artificial selection has proceeded by restricting breeding to young adults, releasing the late part of the life history from natural selection, an approach originally proposed by Edney and Gill<sup>43</sup>. For instance, restriction of the opportunity for reproduction in *Drosophila* to 3–6-day-old flies for 120 generations resulted in a fall in the late but not the early-life fecundity of females in these 'r' lines relative to that of females of 'K' lines where adults of any age could breed<sup>91</sup>. The fall in late-life fecundity was much smaller in crosses between 'r' lines, showing that the genetic basis of selection response differed between them, and that senescence had evolved by accumulation of partially recessive deleterious alleles, which would not be expected on the optimization theory. Further support for mutation accumulation came from the slow appearance of the drop in late fertility. But the *r* and *K* lines were cultured at different densities, so that selection on life-history traits other than late fecundity may also have differed between them and may have accounted for the decline in late fertility of the *r*-line females. In addition, population size was smaller in the *r* lines, so that inbreeding depression may have contributed to the difference from the *K* lines.

The experimental results suggest that ageing in *Drosophila* has evolved in part as a consequence of selection for an optimal life history, and in part as a result of accumulation of predominantly late-acting deleterious mutations. But the magnitude of the contribution from mutation is unknown. Quantification of these effects presents a major challenge for the future.

## Decline and extinction of lineages

If some of the germ lines that led from the primaevial soup had not been, to a first approximation, immortal then extant organisms would not exist. But germ lines can deteriorate and cause extinction of lineages. Long-term cultures of unicellular organisms in small, asexual populations have shown a steady decline in the rate of division, and cultures sometimes went extinct, no matter how carefully they were kept. Senescence could be prevented or reduced either by retaining binary fission but keeping the population size large, or by allowing sexual reproduction<sup>5</sup>. This kind of clonal ageing has also been found in multicellular organisms reproducing asexually by apomixis, fission or budding<sup>5</sup>, although the data for vegetative reproduction are mixed.

Senescence of clones is probably caused by the accumulation of deleterious mutations. The simplest case is where the clone is propagated through a single, randomly chosen cell. There is then no selection, and an inevitable decline in viability and

fertility. Even if a moderately large number (*N*) of individuals is used, mutations are likely to be fixed by chance if they reduce fitness by less than  $s \approx 1/N$ . Moreover, with strictly asexual reproduction, mutations will inevitably accumulate. Even in a large population, very few individuals may be free of any deleterious mutations; if this fittest class fails to leave descendants, it can never be recovered, and the mean fitness of the population will decline irreversibly, in a process known as 'Muller's ratchet'<sup>92,93</sup>. In multicellular organisms that reproduce asexually through a single-cell stage, the effects of Muller's ratchet will be exacerbated, both because population sizes tend to be smaller than for protozoans, and because genomes tend to be large and hence to accumulate more mutations per genome. Sex may therefore have been a necessary prelude to the evolution of multicellular organisms<sup>5,6</sup>. Immortal germ lines require, first, efficient systems for the endogenous repair of biochemical damage and, second, large population sizes and sexual reproduction, which allow elimination of the unfit by natural selection (a kind of exogenous repair<sup>5</sup>). Persistent asexual lineages of organisms with large genomes would therefore present something of a paradox, and they do not seem to occur, at least in animals. Ostensibly asexual lineages of fish and salamanders have been shown by analysis of their mitochondrial DNA to be of some antiquity (4–5 million and 100,000 years, respectively). But neither is truly asexual. Analysis of the nuclear DNA of the mole salamander *Ambystoma* has shown that there has been crossing with males of related sexual species<sup>94,95</sup>. In the fish *Poeciliopsis* males of a sexual species contribute to every 'unisexual' offspring an expressed haploid genome that is then discarded at meiosis<sup>96</sup>. The effects of Muller's ratchet in the maternal complement in these hybrid offspring may therefore be masked to some extent by wild-type alleles in the paternal complement.

Clonal senescence has been thought to bear on the mechanisms of ageing in Metazoan individuals. But there is little direct connection. To the extent that a multicellular soma is not maintained by selection between cells, accumulation of somatic mutations is inevitable. Nevertheless, the empirical evidence is that these contribute little to individual ageing<sup>4</sup>. The loss of division potential with age in cultures of mammalian cells is caused by specific mechanisms preventing proliferation in differentiated cells<sup>97,98</sup>, rather than by accumulation of mutations.

## Soma and germ line

Several authors have suggested that ageing should be confined to organisms with "... a clear distinction between soma and germ plasm"<sup>4,10</sup>. But the logic of this point of view is not compelling, and it is not supported by the data<sup>99</sup>. Consider a multicellular organism in which randomly chosen cells dedifferentiate and divide to produce germ cells until the organism has completely turned into gametes, and where there is therefore no distinction between soma and germ line. Will this creature, as an individual, age? As gamete-production proceeds, it will have a non-zero probability of death or of becoming unable to produce gametes. The intensity of selection on survival and rate of gamete-production will therefore decline with age and ageing will evolve. The same argument would apply to a unicellular organism that produces, by asymmetrical cell division, a series of daughter cells that become gametes, before itself dividing to produce gametes. The critical requirement for the evolution of ageing is that there be a distinction between a parent individual and the smaller offspring for which it provides. If the organism breeds by dividing equally into identical offspring, then the distinction between parent and offspring disappears, the intensity of selection on survival and reproduction will remain constant and individual ageing is not expected to evolve.

## Prospects

The polygenic nature of ageing contains a baleful message for gerontologists; it is most unlikely that engineering of a few genes



or intervention in a handful of physiological pathways will prevent the process from occurring. But an evolutionary decrease in the rate of ageing in humans is likely to occur over the next few generations. Senescence in industrialized human societies has become so apparent because of the removal of most extrinsic causes of death and lowered fertility; in the circumstances in which the life history evolved, these would have predominated, and few individuals would have lived long enough to show evidence of ageing. Industrialization has not only uncovered ageing, but has also been accompanied by an increase in the average age of reproduction, and economic factors may mean that this trend continues. Increasing the age at breeding leads to selection for reduced senescence, and both evolutionary theories specify that survival and fertility later in the lifespan will in consequence increase over the generations to come. But the two theories carry very different implications for the nature of the correlated responses. If mutation-accumulation is the

predominant evolutionary cause of ageing in humans, then reduced ageing could evolve by purging of late-acting mutations, with little immediate cost to survival and fertility in earlier life. If, by contrast, ageing has evolved as part of an optimal life history, then there may be a direct cost, for instance in the form of a reduction in survival to maturity, delayed maturity or lower fertility in early adulthood. These predictions assume a constant environment. But as well as removing hazard, industrialization has been accompanied by an improvement in nutrition which may, if it persists, reduce or abolish any early-life costs of the late-life improvement. Unfortunately the authors and contemporary readers of this review are unlikely to see the results of this experiment. Work with other organisms should yield some insight into the likely outcome. □

L. Partridge and N. H. Barton are in the ICAPB, Division of Biological Sciences, Edinburgh University, West Mains Road, Edinburgh EH9 3JT, UK.

1. Comfort, A. *The Biology of Senescence* 3rd edn (Churchill Livingstone, Edinburgh and London, 1979).
2. Charlesworth, B. *Evolution in Age-Structured Populations* (Cambridge Univ. Press, Cambridge, 1980).
3. Finch, C. E. *Longevity, Senescence and the Genome* (University of Chicago Press, Chicago, 1990).
4. Rose, M. R. *Evolutionary Biology of Ageing* (Oxford Univ. Press, Oxford, 1991).
5. Bell, G. *Sex and Death in the Protozoa* (Cambridge Univ. Press, Cambridge, 1988).
6. Kondrashov, A. S. *Nature* **336**, 435-441 (1988).
7. Nesse, R. M. *Exp. Geront.* **23**, 445-453 (1988).
8. Partridge, L. in *Lifetime Reproduction in Birds* (ed. Newton, I.) 421-440 (Academic, London, 1989).
9. Promislow, D. E. L. *Evolution* **45**, 1869-1887 (1991).
10. Williams, G. C. *Evolution* **11**, 398-411 (1957).
11. Bernstein, C. & Bernstein, H. *Ageing, Sex and DNA Repair* (Academic, San Diego, 1991).
12. Kirkwood, T. B. L. *Mutation Res.* Pilot issue 7-13 (1988).
13. Orgel, L. E. *Proc. natn. Acad. Sci. U.S.A.* **49**, 517-521.
14. Tyndale-Biscoe, M. *Bull. Entomol. Res.* **74**, 341-377 (1984).
15. Kirkwood, T. B. L. in *Physiological Ecology: an Evolutionary Approach to Resource Use* (eds Townsend, C. R. & Calow, P.) Blackwell Scientific, Oxford, 1981).
16. Medawar, P. B. *Modern Quarterly* **2**, 30-49 (1946) (reprinted in *The Uniqueness of the Individual*, Methuen, London).
17. Medawar, P. B. *An Unsolved Problem in Biology* (Lewis, London, 1952).
18. Partridge, L. & Fowler, K. *Evolution* **46**, 76-91 (1992).
19. Roper, C., Pignatelli, P. & Partridge, L. *Evolution* (in the press).
20. Turner, J. R. G. *A. Rev. Ecol. Syst.* **12**, 99-122 (1981).
21. Bulmer, M. G. *Genetics* **129**, 897-907 (1991).
22. Maynard Smith, J. *A. Rev. Ecol. Syst.* **9**, 31-56 (1978).
23. Parker, G. A. & Maynard Smith, J. *Nature* **348**, 27-33 (1990).
24. Stearns, S. C. *Q. Rev. Biol.* **51**, 3-47 (1976).
25. Partridge, L. & Harvey, P. H. *Science* **241**, 1449-1455 (1988).
26. Lessells, C. M. in *Behavioural Ecology: an Evolutionary Approach* 3rd edn (eds Krebs, J. R. & Davies, N. B.) 32-68 (Blackwell Scientific, Oxford, 1991).
27. Stearns, S. C. *The Evolution of Life Histories* (Oxford Univ. Press, Oxford, 1992).
28. Pease, C. M. & Bull, J. J. *J. evol. Biol.* **1**, 293-303 (1988).
29. Charnov, E. L. *Heredity* **62**, 97-106 (1989).
30. Charlesworth, B. *Evolution* **44**, 520-538 (1990).
31. Charnov, E. L. *J. evol. Biol.* **3**, 139-144 (1990).
32. Hamilton, W. D. *J. theor. Biol.* **12**, 12-45 (1966).
33. Williams, G. C. *Am. Nat.* **100**, 687-690 (1966).
34. Partridge, L. & Harvey, P. *Nature* **316**, 20-21 (1985).
35. Reznick, D. *Oikos* **44**, 257-267 (1985).
36. Bell, G. & Koufopanou, V. *Oxf. Surv. evol. Biol.* **3**, 83-131 (1986).
37. Partridge, L. & Sibly, R. *Phil. Trans. R. Soc. B332*, 3-13 (1991).
38. Van Voorhees, W. A. *Nature* **360**, 456-458 (1992).
39. Kirkwood, T. B. L. & Rose, M. R. *Phil. Trans. R. Soc. B332*, 15-24 (1991).
40. Rose, M. R. *Heredity* **48**, 63-78 (1982).
41. Rose, M. R. *Theor. Pop. Biol.* **28**, 342-358 (1985).
42. Rose, M. R., Service, P. & Hutchinson, E. W. in *Genetic Constraints on Adaptive Evolution* (ed. Loeschcke, V.) 91-105 (Springer, Berlin, 1987).
43. Edney, E. B. & Gill, R. W. *Nature* **220**, 281-282 (1968).
44. Kimura, M. & Maruyama, T. *Genetics* **54**, 1337-1351 (1966).
45. Haldane, J. B. S. *Am. Nat.* **71**, 337-349 (1937).
46. Muller, H. J. *Am. J. hum. Genet.* **2**, 111-176.
47. Crow, J. F. & Kimura, M. *An Introduction to Population Genetics Theory* (Harper & Row, New York, 1970).
48. Fisher, R. A. *Am. Nat.* **62**, 571-574 (1928).
49. Charlesworth, B. *Nature* **278**, 848-849 (1979).
50. Maynard Smith, J. *Am. Nat.* **110**, 331-338 (1976).
51. Stenseth, N. C. & Maynard Smith, J. *Evolution* **38**, 870-880 (1984).
52. Pomeroy, D. *Biol. J. Linn. Soc.* **40**, 53-65 (1990).
53. Austad, S. N. & Fischer, K. E. *J. Geront.* **46**, 47-53 (1991).
54. Crow, J. F. & Simmons, M. J. in *The Genetics and Biology of Drosophila* Vol. 3c (eds Ashburner, M., Carson, H. L. & Thompson, J. N.) 1-235 (Academic, New York, 1983).
55. Kondrashov, A. S. & Turelli, M. *Genetics* **132**, 603-618 (1992).
56. Houle, D. *Genetics* **130**, 195-204 (1992).
57. Luckinbill, L. S., Graves, J. L., Reed, A. H. & Koetsawang, S. *Heredity* **60**, 367-374 (1988).
58. Hutchinson, E. W. & Rose, M. R. *Genetics* **127**, 719-727 (1991).
59. Hutchinson, E. W., Shaw, A. J. & Rose, M. R. *Genetics* **127**, 729-737 (1991).
60. Barton, N. H. & Turelli, M. *A. Rev. Genet.* **23**, 337-370 (1989).
61. Rose, M. & Charlesworth, B. *Nature* **287**, 141-142 (1980).
62. Hughes, D. M. & Clark, A. G. *Evolution* **42**, 1309-1320 (1988).
63. Tucic, N., Cvetkovic, D. & Milanovic, D. *Heredity* **60**, 55-60 (1988).
64. Scheiner, S. M., Caplan, R. L. & Lyman, R. F. *Evol. Ecol.* **3**, 51-63 (1989).
65. Giesel, J. T. *Am. Nat.* **128**, 593-603 (1986).
66. Shaw, R. G. *Evolution* **41**, 812-826 (1987).
67. Rose, M. R. *Amer. Nat.* **123**, 565-569 (1984).
68. Rose, M. R. *Evolution* **38**, 1004-1010 (1984).
69. Luckinbill, L. S., Arking, R., Clare, M. G., Cirocco, W. C. & Buck, S. A. *Evolution* **38**, 996-1003 (1984).
70. Charlesworth, B. in *Evolutionary Ecology* (ed. B. Shorrocks) 117-133 (1984).
71. Houle, D., Hoffmaster, D. K., Assimacopoulos, S. & Charlesworth, B. *Nature* **359**, 58-60 (1992).
72. Service, P. M., Hutchinson, E. W. & Rose, M. R. *Evolution* **42**, 708-716 (1988).
73. Lande, R. *Genetics* **94**, 203-215 (1980).
74. Wagner, G. *Genetics* **122**, 223-234 (1987).
75. Turelli, M. *Genetics* **111**, 165-195 (1985).
76. Houle, D. *Evolution* **45**, 630-648 (1991).
77. Barton, N. H. *Genetics* **124**, 773-782 (1990).
78. Giesel, J. T., Murphy, P. A. & Manlove, M. N. *Am. Nat.* **119**, 464-479 (1982).
79. Clark, A. G. *Am. Nat.* **129**, 932-940 (1987).
80. Service, P. M. & Rose, M. R. *Evolution* **39**, 943-945 (1985).
81. Maynard Smith, J. *J. exp. Biol.* **35**, 832-842 (1958).
82. Lamb, M. J. *J. Insect Physiol.* **10**, 487-497 (1964).
83. Partridge, L., Green, A. & Fowler, K. *J. Insect Physiol.* **33**, 745-749 (1987).
84. Fowler, K. & Partridge, L. *Nature* **338**, 760-761 (1989).
85. Partridge, L. & Andrews, R. *J. Insect Physiol.* **31**, 393-395 (1985).
86. Service, P. M. *J. Insect Physiol.* **35**, 447-452 (1989).
87. Luckinbill, L. S., Graves, J. L., Tomkiw, A. & Sowirka, O. *Evol. Ecol.* **2**, 85-94 (1988).
88. Barrett, S. C. H., Morgan, M. T. & Husband, B. C. *Evolution* **43**, 1398-1416 (1989).
89. Charlesworth, D., Morgan, M. T. & Charlesworth, B. *Evolution* **44**, 1469-1489 (1990).
90. Orr, H. A. *Proc. natn. Acad. Sci. U.S.A.* **88**, 11413-11415 (1991).
91. Mueller, L. D. *Proc. natn. Acad. Sci. U.S.A.* **64**, 974-1977 (1987).
92. Haigh, J. *Theor. Pop. Biol.* **14**, 251-267 (1978).
93. Bell, G. *J. evol. Biol.* **1**, 67-82 (1988).
94. Spolsky, C. M., Phillips, C. A. & Uzzell, T. *Nature* **356**, 706-708 (1992).
95. Hedges, S. B., Bogart, J. P. & Maxson, L. R. *Nature* **356**, 708-710 (1992).
96. Quattro, J. M., Avise, J. C. & Vrijenhoek, R. C. *Proc. natn. Acad. Sci. U.S.A.* **89**, 348-352 (1992).
97. Goldstein, S. *Science* **249**, 1129-1133 (1990).
98. Raff, M. *Nature* **356**, 397-400 (1992).
99. Martinez, D. E. & Levinson, J. S. *Proc. natn. Acad. Sci. U.S.A.* **89**, 9920-9923 (1992).

ACKNOWLEDGEMENTS. We thank B. Charlesworth, T. Chapman, K. Dawson, K. S. Gale, P. Harvey, A. Kondrashov, J. Maynard Smith, M. J. Morgan, M. Slatkin and M. Turelli for helpful comments and C. Roper for providing the data for Fig. 1. Our work was supported by grants from the NERC and SERC and by the Darwin Trust of Edinburgh.

# Glial growth factors are alternatively spliced erbB2 ligands expressed in the nervous system

Mark A. Marchionni\*, Andrew D. J. Goodearl†, Maio Su Chen\*, Olivia Bermingham-McDonogh\*, Cassandra Kirk\*, Marvin Hendricks\*, Frank Danehy\*, Don Misumi\*, Judith Sudhalter\*, Kazumi Kobayashi\*‡, Diana Wroblewski\*, Catherine Lynch\*, Mark Baldassare§, Ian Hiles†, John B. Davis†‡, J. Justin Hsuan†, Nicholas F. Totty†, Masayuki Otsu†, Robert N. McBurney\*, Michael D. Waterfield†, Paul Stroobant† & David Gwynne\*

\* Cambridge NeuroScience Inc., One Kendall Square, Building 700, Cambridge, Massachusetts 02139, USA

† Ludwig Institute for Cancer Research, Courtauld Building, 91 Riding House Street, London W1P 8BT, UK

§ Columbia University, Howard Hughes Medical Institute, 722 168th Street, New York, New York 10032, USA

‡ Present addresses: Biogen Inc., 14 Cambridge Center, Cambridge, Massachusetts 02142, USA (K.K.); Harvard Medical School, Neurobiology Department, 25 Shattuck Street, Building 52, Boston, Massachusetts 02115, USA (A.D.J.G.); The Salk Institute for Biological Studies, 10010 North Torrey Pines Road, La Jolla, California 92037, USA (J.B.D.).

**Glial growth factors, proteins that are mitogenic for Schwann cells, and several ligands for the p185<sup>erbB2</sup> receptor, are products of the same gene. Alternative splicing of the messenger RNA generates an array of putative membrane-attached, intracellular and secreted signalling proteins, at least some of which are expressed in the developing spinal cord and brain. These factors are probably important in the development and regeneration of the nervous system.**

SCHWANN cells have important roles in the development, function and regeneration of peripheral nerves<sup>1-4</sup>. During embryogenesis Schwann cell progenitors migrate from the neural crest<sup>5</sup>, then proliferate along tracts to be occupied by peripheral axons where they support axon extension. The development of peripheral axons seems to rely on Schwann cells, particularly for the formation of the basal lamina<sup>6</sup> and as a major source of trophic factors<sup>7,8</sup>, which regulate neurite outgrowth and neuronal survival<sup>1-4,9-11</sup>. During later stages of development Schwann cells produce the myelin sheath<sup>12</sup>.

Adult Schwann cells normally are quiescent, but retain the capacity for cell division and can dedifferentiate to a highly proliferative state after nerve injury<sup>13</sup>. Myelin breakdown products<sup>13</sup>, inflammation<sup>14,15</sup> and loss of axonal contacts<sup>16,17</sup> provide several potential signal sources to promote an early burst of Schwann cell proliferation following nerve injury. Later, during axonal regrowth a second burst of Schwann cell proliferation precedes remyelination<sup>16</sup>.

Many trophic factors can influence axonal growth and neuronal survival during nervous system development and in the regeneration of peripheral nerves<sup>3,18-20</sup>. As Schwann cells can produce a number of these factors, it is important to understand the signals that regulate Schwann cell proliferation, migration and differentiation. Studies on purified rat sciatic nerve Schwann cells have shown that proliferation *in vitro* depends on mitogenic factors. Bovine brain extracts and bovine pituitary glands have been identified<sup>21</sup> as enriched sources of a mitogenic activity. A mitogen of *M<sub>r</sub>* 31,000 (31K) named glial growth factor (GGF) was purified from bovine pituitary<sup>22</sup>. Schwann cell mitogenic activity from bovine pituitary extracts can be resolved into three activities with different molecular masses: GGF-I (34K), GGF-II (59K) and GGF-III (45K) (A.D.J.G. *et al.*, manuscript submitted). Protein sequence derived from purified GGFs (A.D.J.G., manuscript in preparation) has enabled us to isolate and characterize several complementary

DNA clones encoding proteins with Schwann cell mitogenic activity.

Our analysis of both cDNA and genomic sequences reveals that the single GGF gene encodes multiple Schwann cell mitogens derived from alternatively spliced mRNAs. Interestingly, products of this gene are also specific activators of the p185<sup>erbB2</sup> receptor tyrosine kinase, including the structurally related, previously described heregulins<sup>23</sup> and neu differentiation factor<sup>24</sup>. Moreover, as revealed by *in situ* hybridization analysis, the GGFs appear to be expressed in the nervous system during embryonic development. We therefore conclude that the GGFs are alternatively spliced, putative p185<sup>erbB2</sup> ligands produced in the embryonic nervous system. GGFs produced in peripheral nerve probably play a significant part in regulating the functions of Schwann cells during development and following nerve injury.

## Sequence of glial growth factors I, II and III

GGF-I, GGF-II and GGF-III were purified from bovine pituitary glands using methods described elsewhere (A.D.J.G. *et al.*, manuscript submitted). Novel peptide sequences obtained after enzymatic digestion are shown in Table 1. A closely related peptide sequence was found to be present in all three GGF molecules, namely GGF-I peptides 2 and 7, GGF-II peptide 12 and GGF-III peptide 13. As these peptides are derived from three separate protein preparations containing GGF activity, we focused on this peptide sequence for the molecular cloning of GGF DNA sequences.

## Cloning and structural analysis

We used degenerate oligonucleotide probes designed from the peptide sequence common to all three peaks of GGF activity to screen a number of cDNA and genomic libraries and initially isolated a single bovine genomic DNA. Characterization of this clone ( $\lambda$ GGFGB1) allowed the isolation of multiple cDNAs



derived from several mammalian tissue sources. Beginning with the bovine source, where the purification of GGF activities from pituitary suggested an apparent tissue source for cDNA cloning, we analysed mRNA prepared separately from anterior lobes, posterior lobes, and hypothalami. These reverse transcription-amplification (RT-PCR) analyses indicated that the posterior lobe of the pituitary was the most enriched bovine neuroendocrine source of mRNA homologous to the genomic sequence determined in AGGF1; consequently, a set of five overlapping PCR products were cloned and sequenced. Analysis of the bovine mRNA and genomic sequences (not shown) showed that the GGF transcripts were alternatively spliced products of a single bovine gene. Collectively, these sequences were found to encode twelve of the pituitary GGF peptides.

To obtain full-length cDNA clones, we screened libraries prepared from bovine posterior pituitary, human fetal brain, human brain stem and rat pituitary. We identified two bovine posterior pituitary cDNAs (GGFBPP4 and GGFBPP5), three human cDNAs, including one from fetal brain (GGFHFB1) and two from brain stem (GGFHBS5 and GGFHBS14), and one rat pituitary cDNA (GGFRP3).

The deduced amino-acid sequences of three full-length clones, GGFHBS5, GGFHFB1 and GGFBPP5 are shown in Fig. 1a. The latter two clones are homologous structures derived from human and bovine sources, respectively, and their deduced sequences are 92% identical. Although GGFHBS5 and GGFHFB1 both contain five common coding segments, these clones commence with distinct N-terminal coding segments designated 1 or 2. Transcription of these two mRNAs is probably initiated at two distinct promoters. Untranslated sequences upstream of the coding regions 1 and 2 were found to be unrelated, and attempts to identify transcripts that contain linked 1 and 2 sequences by library screening or through RT-PCR experiments were unsuccessful (data not shown). But as the 5' ends of these transcripts have not been mapped, we cannot rule out the possibility that transcription is initiated from a single site and that the variability is generated through differential splicing alone.

Fifteen bovine peptide sequences (Table 1) have been mapped within these three deduced bovine and human protein sequences. Segment 1 encodes six of the peptides, all of which are unique to GGF-II. In fact, all but one of the new peptide sequences obtained in the analysis of GGF-II have been accounted for within the predicted GGFHBS5 protein sequence. This correlation of the GGFHBS5 clone with GGF-II activity is supported by estimates of the deglycosylated size (45K) of GGF-II (A.D.J.G. *et al.*, manuscript submitted), which is roughly the predicted size of the protein encoded by GGFHBS5.

The deduced protein sequences reveal an extended hydrophobic stretch located in coding segment 1, near the N terminus of GGFHBS5. Positions of the cysteines indicated in Fig. 1a imply similarity to several known superfamilies of proteins. The pattern of six cysteines encoded in segments 6 and 8 matches that of an epidermal growth factor (EGF)-like motif<sup>25</sup>. Coding segments 3 and 4 contain a pair of cysteines and other landmark residues conserved in members of the C-2 immunoglobulin superfamily<sup>26</sup>. Most significantly, an interesting similarity to the human basement membrane heparan sulphate proteoglycan core protein<sup>27</sup> was detected. The peptide sequence ADSGEY, which is located next to the second cysteine of the C-2 immunoglobulin fold in these GGFs, occurs in nine of the twenty-two C-2 repeats found in that basal lamina protein. Finally, juxtaposition of cysteines 3 and 4 (Cys 245 and Cys 248) in segment 1 may imply a kringle-type folding pattern for the N terminus. Alternatively, these cysteine residues could participate in intermolecular disulphide bonds.

While these studies were being completed, five different cDNAs encoding several putative ligands of the p185<sup>erbB2</sup> receptor tyrosine kinase were cloned and characterized<sup>23,24</sup>. Comparison of our GGF clones with these sequences reveals that

TABLE 1 Peptide analysis of bovine glial growth factors I, II and III

| Protein | Peptide number | Peptide sequence  | Protease |
|---------|----------------|-------------------|----------|
| GGF-I   | 2              | ASLADEYEYMXK*     | Trypsin  |
|         | 7              | ASLADEYEYMRK*     | Trypsin  |
|         | 3              | TETSSEYLXLK*      | Trypsin  |
|         | 18             | EYKCLKFKWFKKATVM† | V8       |
| GGF-II  | 1              | VHQVWAAK*         | Trypsin  |
|         | 2              | YIFFMEPEAXSSG     | Trypsin  |
|         | 3              | LGAWGPPAFPVXY     | Trypsin  |
|         | 4              | WVFVIEGK*         | Trypsin  |
|         | 6              | LVLK*             | Trypsin  |
|         | 8              | ASPVSVGSVQELVQR*  | Trypsin  |
|         | 10             | DLLXV             | Trypsin  |
|         | 11             | KVHQVWAAK*†‡      | Lysyl    |
| GGF-III | 12             | KASLADSGEYMXK*†‡  | Lysyl    |
|         | 11             | KSELRIK*†         | Lysyl    |
|         | 12             | KLGNDSANITIV†     | Lysyl    |
|         | 13             | KASLADSGEYMCK*†   | Lysyl    |
|         | 14             | KWFK*†            | Lysyl    |
|         | 15             | KDLSNPSRYLCK*†    | Lysyl    |

Protein sequences were obtained from proteolytic digests of GGF-I, GGF-II and GGF-III, purified as described (A.D.J.G. *et al.*, manuscripts in preparation). The proteases were trypsin, protease V8 (V8), and *Achromobacter lyticus* lysylendopeptidase (Lysyl). N-terminal lysine residues were assumed to be present in lysylendopeptidase digests because of the high specificity of this enzyme.

\* Peptide was completely sequenced.

†‡ Proteins were digested directly, or after reduction, alkylation and desalting (†), or after reduction and purification by gel electrophoresis (‡). Gel-purified GGF-III was digested directly in the excised gel slice (J.H., unpublished results) and GGF-II was electroeluted before digestion. Peptides from these digests were separated by reverse phase or tandem ion exchange-reverse phase HPLC. Peaks detected by a diode array were sequenced by a modified automated pulsed liquid Edman system. Cysteine residues in the GGF-III digests were only identified later by recognition of a phenylthiohydantoin-cysteine derivative (N.T., unpublished results).

the glial growth factors and those ligands, namely the neurotrophin differentiation factor (NDF) and the heregulins, are products of the same gene. Heregulin  $\beta 3$  (ref. 24) is identical to GGFHFB1, but GGFHBS5 is a new structure. In comparing the deduced sequences of the two human GGF cDNA clones isolated from tissue sources (GGFHBS5 and GGFHFB1), we have confirmed the sequence variation noted<sup>23</sup> between  $\alpha$ - and  $\beta$ -heregulins (position 38 in GGFHFB1). This residue is therefore probably polymorphic in the human population.

To understand the structure of specific parts of the GGF gene most relevant to the mRNAs that we have identified, we analysed sequences from two bovine genomic DNAs containing coding segments 1, 3 and 4. DNA sequences extending across these three coding regions have been determined, and these data showed that the coding segments 1, 3 and 4 each represent exons in the bovine genome. Inspection of the intron/exon junction sequences (Fig. 1b) obtained in the analyses of these three exons has shown that the different N termini of GGFHBS5 and GGFHFB1 (and the homologous bovine mRNAs) can be explained readily by alternative splicing. Either exon 1 or coding segment 2 can splice with exon 3 to encode an alanine across the splice junction. Secondly, only exon 3 can splice with exon 4 to encode a glycine across the splice junction and maintain the reading frame. As the same alanine and glycine codons are present at these positions in both human and bovine cDNA sequences and because the human gene, where analysed (data not shown), has the same intron/exon boundaries, we suggest that these conclusions are applicable to the GGF gene structure and to the mRNA splicing patterns in both species.

The GGF gene has a mosaic structure, composed of at least 13 coding segments (representing a minimum number of exons) and at least two functional domains recognized in known

superfamilies as already described. The transcripts start with either coding segment 1 or segment 2; coding segments 1 and 5 are mutually exclusive. With one exception, GGFBBP1, all of the mRNAs contain a core immunoglobulin-like fold and one EGF-like motif (segments 3, 4, 6, 7 or 3, 4, 6, 8). Variations on these two central themes occur in both the 5' and 3' directions (Fig. 2) and contained within the core sequence is the optional segment 5. Interestingly, the 3' end of GGFBBP1 differs significantly from all others: C-terminal coding sequences and the 3' untranslated sequences are transcribed by reading beyond exon 4 to a polyadenylation site in the following intron sequence (Fig. 1c); this transcript encodes a protein that is truncated after the immunoglobulin-like domain. We postulate that GGFBBP1 and/or a related transcript (containing segment 2 as an N-terminal coding segment instead of segment 1) may lack mitogenic activity. By analogy, a splicing variant of the hepatocyte growth factor (HGF) gene produces a truncated protein, which is a natural antagonist of HGF<sup>28</sup>. Possibly this form of GGF could act as a natural antagonist by inhibiting an activity of the receptor, such as ligand binding, dimerization or kinase activation. This suggestion is supported by studies that show that at least some of the activities of GGF (and related)

gene products have been associated with expression of the EGF motif alone<sup>23</sup>.

Although the three purified GGFs are products of one gene, specific assignment of a single mRNA to each protein is made difficult by the large number of alternatively spliced mRNA variants. But as already noted, the predicted sequence and size of the protein encoded by GGFHBS5 make it a likely candidate to encode GGF-II. We have therefore designated the product of this sequence as recombinant human (rh) GGF-II.

### Activities of GGFs on cultured Schwann cells

We transiently expressed several cDNAs in mammalian cells, then assayed both cell lysates and conditioned medium in a Schwann cell proliferation assay. We observed a twofold to fivefold stimulation of proliferation of Schwann cells treated with recombinant materials over background activity present in mock-transfected host cells (Table 2). The mitogenic activity of the GGFHFB1-encoded protein was detected within lysates derived from transfected cells and was not detected in conditioned medium. Similar results to GGFHFB1 were obtained with the bovine cDNA clone GGFBBP5 (data not shown). In view of the sequence identity of GGFHFB1 and heregulin  $\beta$ 3,

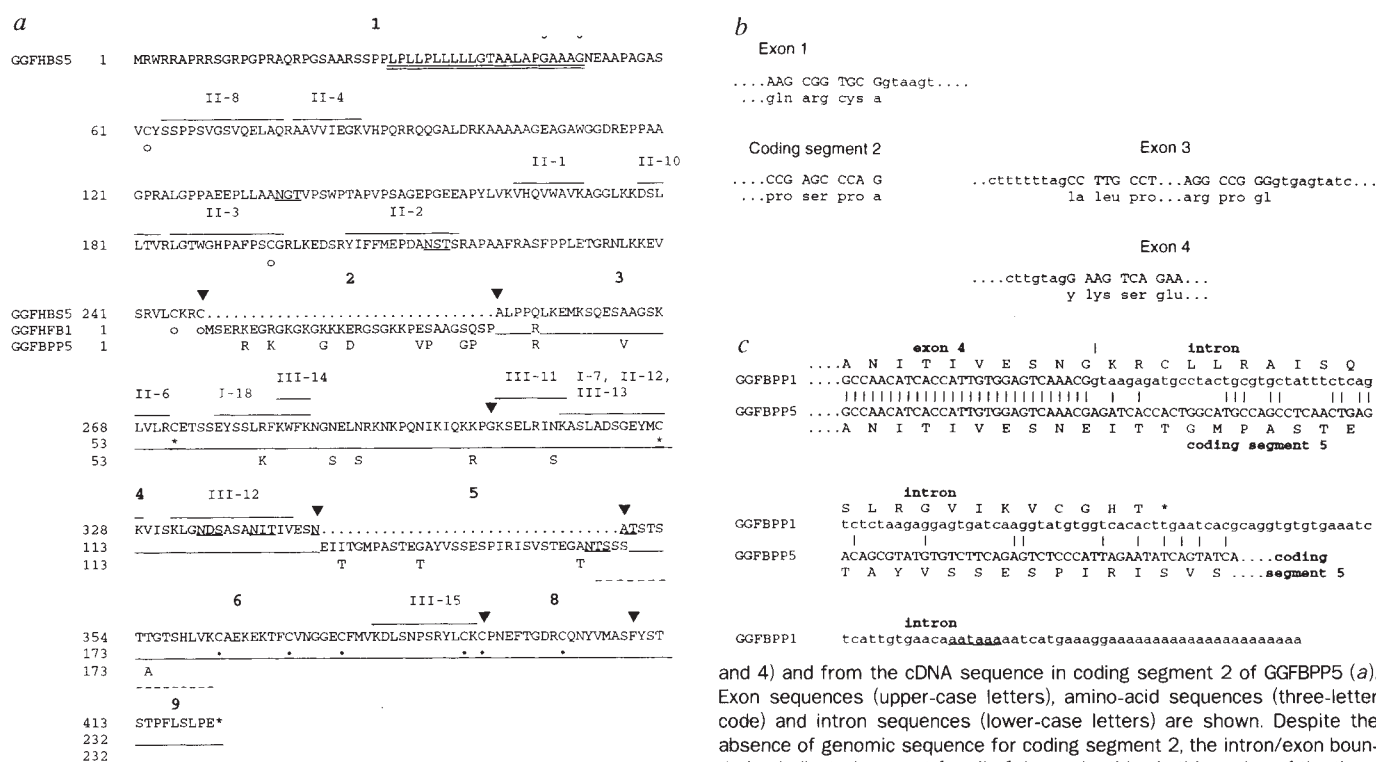


FIG. 1. a, Deduced sequences of three human and bovine glial growth factor clones. Sequences from the human brain stem cDNA clone GGFHBS5 (top), the human fetal brain cDNA clone GGFHFB1 (middle) and the bovine posterior pituitary cDNA clone GGFBBP5 (bottom; only differences from human sequence) are shown. The boundaries (▼) of coding segments (1-6, 8, 9, bold face; see also Fig. 2) are designated, as are the positions (overline) of bovine GGF peptides listed in Table 1. Discrepancies between the deduced amino-acid sequence in this figure and peptide sequences in Table 1 can be ascribed to difficulties in the assignment of some derivatized amino acids, or due to divergence of bovine and human coding sequences. Coding segments absent from GGFHBS5 are shown as dotted lines; segments of identity between GGFHBS5 and GGFHFB1 are shown as a solid line. A hydrophobic stretch (double underline) is indicated near the N terminus. Cysteine residues are indicated that belong to putative kringle-like (○), immunoglobulin-like (\*) and epidermal growth factor-like (●) folds. Sites of potential covalent modification include N-terminal myristylation (V), and N-linked (—) or O-linked (dashed line) carbohydrate attachment. Accession numbers are L12259-61. b, Bovine gene sequences at intron/exon boundaries. Splice junction sequences are shown from three bovine exons (1, 3

and 4) and from the cDNA sequence in coding segment 2 of GGFBBP5 (a). Exon sequences (upper-case letters), amino-acid sequences (three-letter code) and intron sequences (lower-case letters) are shown. Despite the absence of genomic sequence for coding segment 2, the intron/exon boundaries indicated account for all of the nucleotides in this region of the three cDNA clones shown in a. c, GGFBBP1 encodes a truncated protein. Nucleotide and deduced amino-acid sequences of two bovine cDNAs, GGFBBP1 and GGFBBP5, are shown in the areas of exon 4 and coding region 5. The 3' end of GGFBBP1 includes genomic sequence (lower-case letters) from the adjoining intron, which is spliced out of all other transcripts (Fig. 2). A canonical polyadenylation sequence (underlined) is followed by a 22-nucleotide poly (A) tail.

**METHODS.** A set of ten overlapping degenerate probes representing bovine GGF-II peptide 12 were radiolabelled and used to screen a bovine genomic library (Stratagene). One genomic DNA hybridized to two probes. Sequences that encode GGF peptides were extended by PCR amplification of RNA prepared from frozen bovine tissue (Pelfreeze) according to ref. 46. Polyadenylated RNA was selected and converted to cDNA (Perkin Elmer PCR/RNA kit). Amplification was by RACE reaction<sup>47</sup> or as an anchored PCR reaction. Specific amplification products detected on Southern blots with a radiolabelled internal probe were subcloned and sequenced. Fragments from these cDNAs and unique oligonucleotides were radiolabelled and used for probing cDNA libraries constructed in the  $\lambda$  Zap II vector (Stratagene) by standard techniques<sup>48</sup>, or were purchased from Clontech and Stratagene.



TABLE 2 GGF secretion in COS-7 cells

| Transfected DNA                        | Activity in cell lysate |       | Activity in conditioned medium |        | Distribution of total activity (%) |                    |
|----------------------------------------|-------------------------|-------|--------------------------------|--------|------------------------------------|--------------------|
|                                        | 10 $\mu$ l              | Total | 10 $\mu$ l                     | Total  | Cell lysate                        | Conditioned medium |
| <sup>125</sup> I-uridine incorporation |                         |       |                                |        |                                    |                    |
| GGF mock                               | 183 $\pm$ 49            | —     | 50 $\pm$ 7                     | —      |                                    |                    |
| GGFHFB1                                | 391 $\pm$ 98            | 915   | 67 $\pm$ 14                    | ND     | 100                                | —                  |
| GGFHBS5                                | 986 $\pm$ 111           | 9,645 | 182 $\pm$ 33                   | 64,400 | 13                                 | 87                 |
| $\beta$ -Galactosidase units           |                         |       |                                |        |                                    |                    |
| $\beta$ -gal mock                      | 0.05                    | —     | 0.05                           | —      |                                    |                    |
| $\beta$ -gal                           | 0.82                    | 12.3  | 0.05                           | ND     | 100                                | —                  |

Schwann cell mitogenic activity was measured in quadruplicate and the level of [<sup>125</sup>I]uridine incorporation associated with mock-transfected samples (mean plus one standard deviation) was subtracted from the activity (mean minus one standard deviation) in the cell lysate and in the conditioned medium to calculate total GGF activity in each sample. Total volumes were 150  $\mu$ l for lysate samples and 7 ml for conditioned medium samples. Expression from all GGF transfections was judged to be comparable by northern blot analysis of RNA. ND, not detected.  $\beta$ -Galactosidase ( $\beta$ -gal) co-transfections were analysed by a colorimetric assay (Promega), in which 1 unit  $\beta$ -galactosidase activity is represented by one absorbance unit at 410 nm. The value from a mock-transfected sample was subtracted to calculate the distribution of activity. Mitogenic responses of cultured Schwann cells were measured in the presence of 5  $\mu$ M forskolin after an 18–24 h exposure to materials obtained from transfected or mock-transfected cos cells. cDNAs (Fig. 1) were cloned into pCDL-SR $\alpha$ 296 (ref. 50), and COS-7 cells were transfected in 100-mm dishes by the DEAE-dextran method<sup>48</sup>. Cell lysates or conditioned media were collected at 3 or 4 days post-transfection. To prepare lysates, cell monolayers were washed with PBS, scraped from the dishes, and lysed by three freeze–thaw cycles in 150  $\mu$ l 0.25 M Tris–HCl, pH 8. Cell debris was pelleted and the supernatant recovered. Conditioned media samples (7 ml) were collected, then concentrated and buffer exchanged with 10 mM Tris, pH 7.4 using Centrprep-10 and Centricon-10 units as described by the manufacturers (Amicon, Beverly, MA). Rat sciatic nerve Schwann cells were assayed for incorporation of DNA synthesis precursors, as described<sup>36,51</sup>.

the lack of detectable secreted activity was not surprising. In contrast to the other heregulins, biologically active (p185<sup>erbB2</sup> receptor tyrosine kinase activation) heregulin  $\beta$ 3 is not detected in conditioned medium from transiently expressing COS cells<sup>23</sup>. Here, only recombinant human GGF-II (from GGFHBS5) is secreted at significant levels into conditioned medium. To investigate this ability of different recombinant GGFs to be secreted, we determined for the two human cDNA clones the distribution of total expressed GGF (Schwann cell mitogenic) activity between cell lysate and conditioned medium after transient mammalian transfections (Table 2). Most (87%) of the mitogenic activity expressed in GGFHBS5 transfections was present in the conditioned medium. In contrast, we did not detect significant amounts (over mock transfected control samples) of activity in conditioned medium from the GGFHFB1 transfection. Co-transfection and parallel transfection experiments using a non-secreted marker protein ( $\beta$ -galactosidase plasmid) confirmed that GGF activity in conditioned medium results from secretion and not simply from cell lysis or leakage.

As all three recombinant GGFs terminate in the ectodomain, the ability of recombinant GGF-II to be secreted is presumably mediated through the N-terminal hydrophobic stretch encoded by segment 1. It seems less likely that the other major structural differences between GGFHBS5 and GGFHFB1 would influence secretion. As coding segments 2 and 5 are either present (in heregulins  $\alpha$ ,  $\beta$ 1,  $\beta$ 2 and NDF) or absent (from GGFHBS5) in these secreted (and released) recombinant proteins, these regions apparently are not involved in regulating secretion.

### Receptor tyrosine kinase activation

Many peptide growth factors mediate mitogenic responses as a result of activation of receptor tyrosine kinases<sup>29–31</sup>. The *HER-2/neu* oncogene encodes a receptor tyrosine kinase, p185<sup>erbB2</sup> (ref. 32), which is structurally related to the EGF receptor<sup>33</sup> and to erbB3 (ref. 34). Additionally, p185<sup>erbB2</sup> is expressed on the surface of Schwann cells and is upregulated during wallerian degeneration and also by exposure to forskolin<sup>35</sup>, which potentiates the mitogenic response to GGF<sup>36</sup>. As the recombinant GGFs described here and several putative p185<sup>erbB2</sup> ligands are products of the same gene, we reasoned that bovine pituitary GGFs and rhGGF-II might activate a p185 receptor tyrosine kinase (perhaps erbB2) on Schwann cells. Therefore we analysed tyrosine phosphorylation of Schwann cell protein by western blotting (Fig. 3). This demonstrated that Schwann cells treated with either native or recombinant GGFs phosphorylated (on

tyrosine) a protein of 185K. A simple hypothesis that unifies these observations is that GGF-II (or other GGFs) stimulates Schwann cells to divide by activating the tyrosine kinase of p185<sup>erbB2</sup> or a closely related subtype of that receptor.

### Tissue-specific expression

Because glial growth factors could be involved in the development of the peripheral nervous system, we have localized, by preliminary *in situ* hybridization experiments, the sites of GGF transcription in the developing rodent nervous system (Fig. 4). Sections of mouse embryos from E11 to E15 were labelled with probes derived from the common EGF-like region of the GGF gene. The first detectable expression of GGF mRNA was in the nervous system, beginning at embryonic day (E) 10.5 to 11. The highest levels of GGF mRNA are in the spinal cord (Fig. 4a) in the ventral region, coincident with the position of newly differentiated motor neurons. In addition, high levels of GGF mRNA are expressed in dorsal root ganglia (Fig. 4a). The region of the dorsal root ganglion nearest to the spinal cord, the boundary cap region<sup>37</sup> and the peripheral nerve, which contains non-neuronal satellite cells and primitive Schwann cells, do not express GGF mRNA. We therefore suggest that GGF mRNA is localized to neuronal cells present in the ganglion. The pattern of expression in the spinal cord and dorsal root ganglia is maintained at later stages of development (Fig. 4b, c). In E13 embryos, the dorsolateral migration of visceral motor neurons is detected as a small cluster of cells expressing GGF mRNA dorsolateral to the ventral motor column (Fig. 4c). Both somatic and visceral motor neurons seem to express GGF mRNA. GGF mRNA is also expressed in cranial sensory ganglia and in the embryonic brain (Fig. 4d–f). Prominent sites of GGF mRNA expression at E11–E13 include the trigeminal ganglion (Fig. 4e) and other primary motor neuron groups, such as those in the oculomotor nucleus of the midbrain (Fig. 4d) and in the region of newly differentiated neurons in the diencephalon (Fig. 4f). Other regions of the forebrain, including the retina (Fig. 4e) and telencephalon (Fig. 4f) do not express detectable levels of GGF mRNA at this stage of development.

### Discussion

Using peptide sequence information derived from purified bovine glial growth factors, we have isolated several clones that encode Schwann cell mitogenic activities. The group of proteins encoded by these transcripts includes the three described glial growth factors (ref. 22, and A.D.J.G. *et al.*, manuscript

submitted) and several proteins (heregulins and Neu differentiation factor<sup>23,24</sup>) that have various activities on cultured breast tumour cells. The heregulins and NDF are specific activators of the p185<sup>erbB2</sup> receptor<sup>23,24</sup> and rhGGF-II activates the phosphorylation of a Schwann cell protein that is similar in size to p185<sup>erbB2</sup>. Based on several lines of evidence, including the data presented here, we suggest that various p185<sup>erbB2</sup> ligands, such as the GGFs, are important in development, regeneration and tumour formation in the nervous system.

A critical stage in the development of the peripheral nervous system is the projection of axons to their targets, the ensheathment of these axons and their subsequent myelination by Schwann cells. Aspects of the expression of the GGF gene presented here suggest that proteins encoded by this gene have important functions in these events. The two cell types that express GGF mRNA at the highest levels are motor neurons and primary sensory neurons. At later stages of embryonic development, other differentiated neurons in the spinal cord

FIG. 2 The GGF gene and its products. The figure (5' to 3', left to right) represents a composite of our knowledge of the GGF gene based on sequence analysis of 14 different mammalian mRNAs and on partial characterization of the bovine gene structure. Recognizable domains include a hydrophobic region (shown as a deleted H<sub>2</sub>O symbol), a putative kringle-like fold, an immunoglobulin-like fold (Ig), a cluster of carbohydrate attachment sites (glyco), an EGF-like motif (EGF), a membrane-spanning  $\alpha$ -helical segment (TM) and a cytoplasmic extension. Coding sequences are divided into segments by comparing the structures of the known mRNAs. Coding segments 1, 3 and 4 have been characterized as exons (Fig. 1b). Genomic mapping has yet to resolve the order of coding segments 1 and 2, or 9 and 10; these are assigned tentatively as shown. Based on nucleotide sequence and amino-acid sequence comparisons, coding segments 7 and 8 seem to represent a partial internal gene duplication. Their relative order (5' to 3') was determined from the analysis of bovine PCR product GGF BPP2, which probably represents an aberrant splicing pattern. Approximate locations of fifteen GGF peptides are indicated (see Fig. 1 and Table 1 for details). The mRNAs represent nine distinct splicing patterns and are grouped according to the sources used for isolation of cDNA clones (GGFBPP4, 5, GGFHBS5, 14, GGFHFB1, GGF BPP3) and of PCR products (GGFBPP1-3) as described in Fig. 1; four heregulin clones (HRG $\alpha$ , HRG $\beta$ 1-3) were isolated from the human breast line MB-MDA-231 (ref. 23) and Neu differentiation factor (NDF) was cloned from the *ras*-transformed fibroblast cell line RAT-1J (ref. 24). Partial cDNA clones are shown by dotted lines at the 5' end; translation termination codons are indicated by an asterisk.

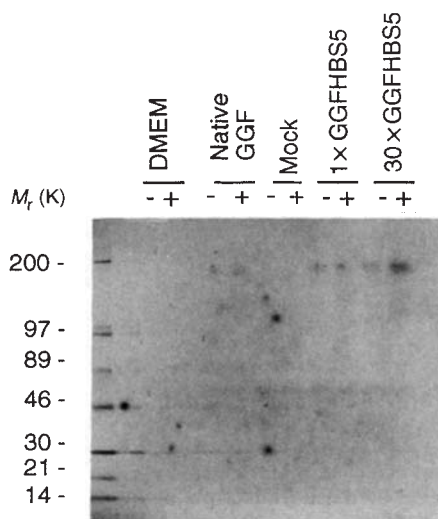
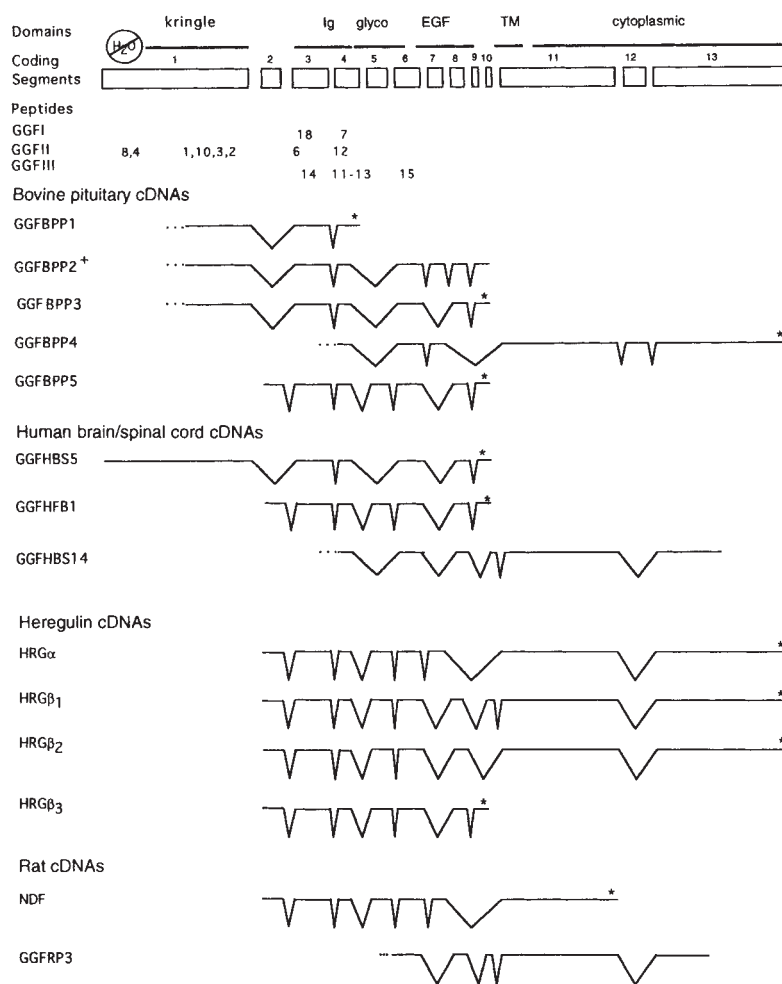


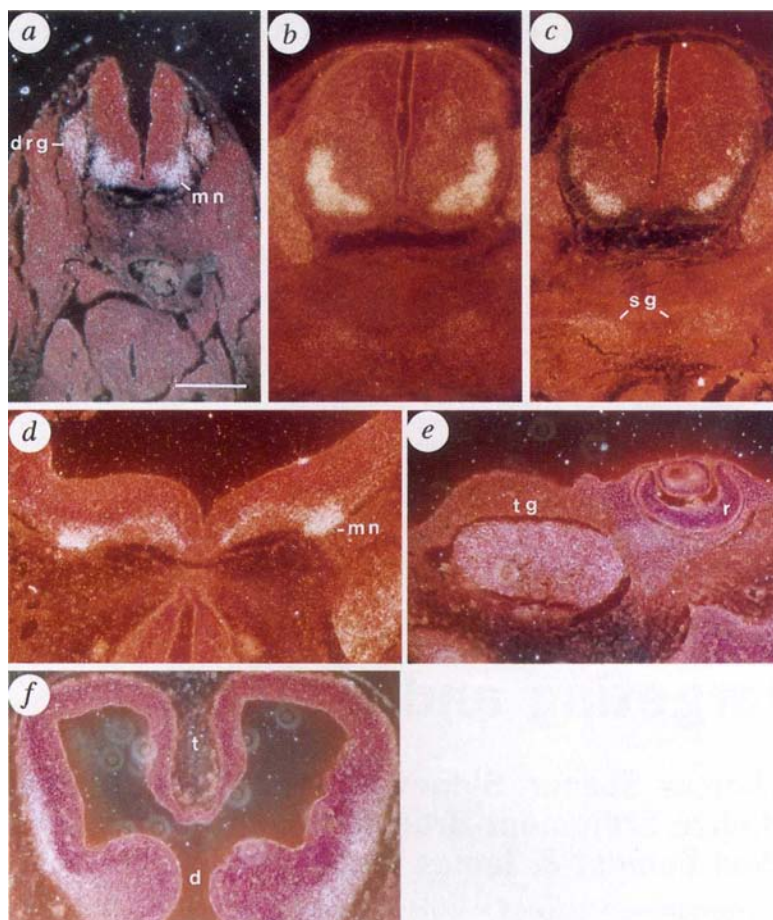
FIG. 3 Tyrosine phosphorylation in Schwann cells treated with recombinant glial growth factors. Rat sciatic nerve Schwann cells were treated for 15 min at 37 °C in the presence (+) or absence (-) of 5  $\mu$ M forskolin with serum-free medium (DMEM), partially purified bovine pituitary GGF I, II and III (native GGF), or with unconcentrated (1 $\times$ ) or 30-fold concentrated (30 $\times$ ) conditioned medium (20  $\mu$ l added to 200  $\mu$ l culture medium) obtained after transient expression (or mock-transfection) of GGFHBS5.

METHODS. Tyrosine phosphorylation was assayed by western blotting as described<sup>23</sup>, using precast 4-12% gradient gels (Novex, San Diego). Membranes were blocked in 4% bovine serum albumin/1% ovalbumin in Tris-buffered saline containing 0.05% Tween-20, pH7.5 (TBST), for a minimum of 2 h before incubating with an <sup>125</sup>I-labelled monoclonal antibody against phosphotyrosine (ICN, Irvine, CA) for 2 h. Blots were washed in TBST and visualized by autoradiography. Protein size standards are myosin (200K), phosphorylase *b* (97K), bovine serum albumin (63K), ovalbumin (46K), carbonic anhydrase (30K), trypsin inhibitor (21K) and lysozyme (14K).



FIG. 4 Localization of GGF mRNA by *in situ* hybridization histochemistry. *a*, Dark-field micrograph showing a section through the caudal region of an E11 embryo. Silver grains are concentrated in the ventral region of the spinal cord, which is consistent with the position of motor neuron cell bodies (mn), and in dorsal root ganglia (drg). Non-neural tissues do not express detectable GGF mRNA. *b* and *c*, Dark-field micrographs of sections through E13 mouse embryos showing high GGF expression in the motor columns of the spinal cord and the dorsal root ganglia. Low levels of hybridization are also detectable in other regions of the spinal cord. In addition, a low level of hybridization is detectable in the region forming sympathetic ganglia (sg). *d*, Dark-field micrograph showing expression of GGF mRNA in the region of the oculomotor neurons (mn) in the ventral midbrain. *e*, Dark-field micrograph showing expression of GGF mRNA in the trigeminal ganglion (tg) in an E11 mouse embryo. The retina (r) does not express detectable GGF mRNA. *f*, Dark-field micrograph of a hybridization of an E11 mouse embryo showing hybridization to cells in the marginal zone of the dien-cephalon (d). Telencephalic regions (t) are negative at this stage. All sections are counterstained with cresyl violet. Scale bar, 200  $\mu$ M in *a-d*; 250  $\mu$ M in *e* and *f*.

**METHODS.** The analysis was done as described in ref. 49, with minor modifications; mice and rats gave essentially similar results. Mouse embryos were fixed by immersion for 2–3 h in 4% paraformaldehyde in 0.1 M phosphate-buffered saline (PBS). Tissue was cryoprotected in PBS containing 20% sucrose at 4 °C overnight, frozen in OTC Tissue-Tek slides, and 10–12  $\mu$ m paraffin sections cut and thaw-mounted onto slides coated with gelatin and poly-L-lysine. A single-stranded RNA probe (specific activity, 0.5–1.0  $\times 10^9$  c.p.m. per  $\mu$ g RNA) was derived from rat cDNA clone GGFRP3 (Fig. 2). The specificity of hybridization was tested using other RNA probes from unrelated transcripts under identical conditions.



also express low levels of GGF mRNA. As axons of central neurons, dorsal root ganglia and sympathetic ganglia<sup>1,2</sup> can induce proliferation of Schwann cells *in vitro*, we suggest that GGF gene products probably support the postnatal proliferation of Schwann cells *in vivo*. The GGFs may also influence other events such as Schwann cell migration, ensheathment and myelination. The recent report that acetylcholine-receptor-inducing activity (ARIA) is associated with the product of the same gene<sup>38</sup> also implicates neuronally produced GGFs in later developmental events in the peripheral nervous system, specifically the formation of the neuromuscular junction.

Many features of regeneration in the peripheral nervous system recapitulate development. Following injury and wallerian degeneration, axons regrow and reinnervate their targets. Here also, various Schwann cell functions essential for regeneration can be regulated by GGF gene products. Several lines of evidence<sup>39,40</sup> point to specific components of the peripheral nerve environment that support regeneration. These components include the basal lamina as well as the trophic factors released by neurons and Schwann cells following injury<sup>3</sup>. Several forms of the GGF (and related) proteins could be involved as normally secreted, or injury-released, activities that regulate the number of Schwann cells and their gene activity. Neuronal membrane-bound forms might function only in close association with target Schwann cells or may be proteolytically clipped and released. A membrane-associated Schwann cell mitogen extracted from neurites of dorsal root ganglia has been described<sup>41</sup>; this mitogen could be a membrane bound form of GGF. Some GGF variants might interact with basal lamina, matrix proteins, or with carrier proteins through specific domains such as the immunoglobulin-like domain.

Because the glial growth factors and specific activators of the p185<sup>erbB2</sup> receptor are encoded by the same gene, we suggest that the formation of tumours derived from Schwann cells, and

perhaps from other glial sources, is regulated by the supply of glial growth factors and by the activity of the p185<sup>erbB2</sup> receptor tyrosine kinase. For example, several tumours of glial cell origin produce GGF-like activities<sup>42</sup>. Moreover, glial tumours express p185<sup>erbB2</sup> (refs 43, 44), a candidate receptor for these proteins that is also associated with adenocarcinoma formation<sup>45</sup>. Some GGFs might also inhibit Schwann cell proliferation, for example the isoform represented in part by the splicing pattern of GGFBPPI.

We propose here a name for the gene we have partly described, and for the protein products (GGFs, NDF, ARIA, heregulins) derived from it. Based on the activation by these proteins of the receptor encoded by *neu*, the neuronal pattern of *in vivo* expression, and the names already given to these proteins (Neu differentiation factor and heregulins), we name the gene neuregulin and refer to the proteins collectively as the neuregulins.

Our findings support the concept that the GGFs and other neuregulins have important functions in the paracrine and autocrine mechanisms that are fundamental to normal development and regeneration in the nervous system and in tumour formation. Analysis of development and regeneration of peripheral nerve tissue using isoform-specific molecular probes should further the understanding of the function of the different forms of GGF. It seems that some of these proteins might have a significant part to play in the treatment of nerve injury, neuropathies and tumours of glial cell origin. □

Received 4 February; accepted 1 March 1993.

1. Bunge, R. P. & Griffin, J. W. in *Diseases of the Nervous System. Clinical Neurobiology* 87–100 (Saunders, Philadelphia, 1992).
2. Jessen, K. R. & Mirsky, R. *Curr. Opin. Neurobiol.* **2**, 575–581 (1992).
3. Fawcett, J. W. & Keynes, R. A. *Rev. Neurosci.* **13**, 43–60 (1990).
4. Eccleston, P. A. *Exp. Cell Res.* **199**, 1–9 (1992).
5. LeDouarin, N. *The Neural Crest* (Cambridge Univ. Press, New York, 1982).

6. Bunge, R. P., Bunge, M. B. & Eldridge, C. F. *A. Rev. Neurosci.* **9**, 305–328 (1986).
7. Muir, D., Gennrich, C., Varon, S. & Manthorpe, M. *Neurochem. Res.* **14**, 1003–1012 (1989).
8. Acheson, A., Barker, P. A., Alderson, R. F., Miller, F. D. & Murphy, R. A. *Neuron* **7**, 265–275 (1991).
9. Baehr, M. & Bunge, R. P. *Expl. Neurol.* **106**, 27–40 (1989).
10. Matsuoka, I., Meyer, M. & Thoenen, H. *J. Neurosci.* **11**, 3165–3177 (1991).
11. Bhattacharyya, A. *et al.* *J. Neurobiol.* **23**, 451–466 (1992).
12. Aguayo, A. J. *et al.* *Brain Res.* **104**, 1–20 (1976).
13. Salzer, J. L., Williams, A. K., Glaser, L. & Bunge, R. P. *J. Cell Biol.* **84**, 753–766 (1980).
14. Baichwal, R. R., Bigbee, J. W. & DeVries, G. W. *Proc. natn. Acad. Sci. U.S.A.* **85**, 1701–1705 (1988).
15. Lu, X. & Richardson, P. M. *J. Neurosci.* **11**, 972–978 (1992).
16. Pellegrino, R. G., Politis, M. J., Ritchie, J. M. & Spencer, P. S. *J. Neurocytol.* **15**, 17–28 (1986).
17. Salzer, J. L. & Bunge, R. P. *J. Cell Biol.* **84**, 739–752 (1980).
18. Brookes, J. P. *Science* **225**, 1280–1287 (1984).
19. Sandrock, A. W. Jr & Matthew, W. D. *Brain Res.* **425**, 360–363 (1987).
20. Sendtner, M., Stöckli, K. A. & Thoenen, H. *J. Cell Biol.* **118**, 139–148 (1992).
21. Raff, M. C., Abney, E., Brookes, J. P. & Hornby-Smith, A. *Cell* **15**, 813–822 (1978).
22. Lemke, G. E. & Brookes, J. P. *J. Neurosci.* **4**, 75–83 (1984).
23. Holmes, W. E. *et al.* *Science* **256**, 1205–1210 (1992).
24. Wen, D. *et al.* *Cell* **69**, 559–572 (1992).
25. Doolittle, R. F., Feng, D. F. & Johnson, M. S. *Nature* **307**, 558–560 (1984).
26. Williams, A. F. & Barclay, A. N. *Rev. Immun.* **6**, 381–405 (1988).
27. Kallunki, P. & Tryggevason, K. *J. Cell Biol.* **116**, 559–571 (1992).
28. Chan, A. M.-L. *et al.* *Science* **254**, 1382–1385 (1991).
29. Yarden, Y. & Ullrich, A. *Rev. Biochem.* **57**, 443–478 (1988).
30. Ullrich, A. & Schlessinger, J. *Cell* **61**, 203–212 (1990).
31. Schlessinger, J. & Ullrich, A. *Neuron* **9**, 383–391 (1992).
32. Schechter, A. L. *et al.* *Nature* **312**, 513–516 (1984).
33. Ullrich, A. *et al.* *Nature* **309**, 418–425 (1984).
34. Kraus, M. H., Issing, W., Miki, T. & Popescu, N. C. *Proc. natn. Acad. Sci. U.S.A.* **86**, 9193–9197 (1986).
35. Cohen, J. A., Yachnis, A. T., Arai, M., Davis, J. G. & Scherer, S. S. *J. Neurosci. Res.* **31**, 622–634 (1992).
36. Davis, J. B. & Stroobant, P. *J. Cell Biol.* **110**, 1353–1360 (1990).
37. Altman, J. & Bayer, S. A. in *Advances in Anatomy Embryology and Cell Biology* Vol. 85 *The Development of the Rat Spinal Cord* (Springer, Berlin, 1984).
38. Falls, D. L., Rosen, K. M., Corfas, G., Lane, W. S. & Fischbach, G. D. *Cell* **72**, 801–815 (1993).
39. Schwab, M. E. & Thoenen, H. *J. Neurosci.* **5**, 2415–2423 (1985).
40. Aguayo, A. J. in *Synaptic Plasticity* 457–484 (Guilford, New York, 1985).
41. Nordlund, M., Hong, D.-M., Fei, X. & Ratner, N. *Glia* **5**, 182–192 (1992).
42. Brookes, J. P., Breakefield, X. O. & Martuza, R. L. *Ann. Neurol.* **20**, 317–321 (1986).
43. Perantoni, A. O., Rice, J. M., Reed, C. D., Watatani, M. & Wenk, M. L. *Proc. natn. Acad. Sci. U.S.A.* **84**, 6317–6321 (1987).
44. Nikitin, A. Y., Ballering, L. A. P., Lyons, J. & Rajewsky, M. F. *Proc. natn. Acad. Sci. U.S.A.* **88**, 9939–9943 (1991).
45. Lupa, R., Dickson, R. B. & Lippman, M. E. *J. Ster. Biochem. molec. Biol.* **43**, 229–236 (1992).
46. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. *Biochemistry* **18**, 5294–5299 (1979).
47. Frohman, M., Dush, M. & Martin, G. *Proc. natn. Acad. Sci. U.S.A.* **85**, 8998 (1988).
48. Sambrook, J., Fritsch, E. F. & Maniatis, T. in *Molecular Cloning. A Laboratory Manual 2nd edn* (Cold Spring Harbor Laboratory press, Cold Spring harbor, New York, 1989).
49. Wilkinson, D. G., Bailes, J. A. & McMahon, A. P. *Cell* **50**, 79–88 (1987).
50. Takebe, Y. *et al.* *Molec. cell. Biol.* **8**, 466–472 (1988).
51. Brookes, J. P., Fields, K. L. & Raff, M. C. *Brain Res.* **165**, 105–118 (1979).

ACKNOWLEDGEMENTS. We thank M. Alfano, G. Gao, J. Lucas, C. Nolan, M. Rossi and O. Truong for technical assistance; G. Fischbach for sharing unpublished results. D. Anderson, J. Brookes, A. Davies, H. Goodman, T. Jessell, R. Kamen, S. J. Lee, M. Noble and A. Server for discussion; L. Guarente for comments on the manuscript; and M. A. Carr for preparing it.

# SNAP receptors implicated in vesicle targeting and fusion

Thomas Söllner, Sidney W. Whiteheart, Michael Brunner, Hediye Erdjument-Bromage, Scott Geromanos, Paul Tempst & James E. Rothman

Rockefeller Research Laboratory, Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, New York, New York 10021, USA

**The *N*-ethylmaleimide-sensitive fusion protein (NSF) and the soluble NSF attachment proteins (SNAPs) appear to be essential components of the intracellular membrane fusion apparatus. An affinity purification procedure based on the natural binding of these proteins to their targets was used to isolate SNAP receptors (SNAREs) from bovine brain. Remarkably, the four principal proteins isolated were all proteins associated with the synapse, with one type located in the synaptic vesicle and another in the plasma membrane, suggesting a simple mechanism for vesicle docking. The existence of numerous SNARE-related proteins, each apparently specific for a single kind of vesicle or target membrane, indicates that NSF and SNAPs may be universal components of a vesicle fusion apparatus common to both constitutive and regulated fusion (including neurotransmitter release), in which the SNAREs may help to ensure vesicle-to-target specificity.**

TRANSACTIONS among the membrane-bound compartments in the cytoplasm of eukaryotic cells are generally executed by transport vesicles which bud from one membrane and fuse selectively with another<sup>1</sup>. The mechanism by which each vesicle chooses its target is currently unknown, but must embody the essence of compartmental specificity. Choice of target is implicit in all vesicular fusion events, from the ubiquitous steps in the constitutive secretory and endocytotic pathways (such as the fusion of a vesicle from the endoplasmic reticulum with the Golgi) to specialized and tightly regulated forms of exocytosis (such as the triggered fusion of a synaptic vesicle containing neurotransmitter with the axonal membrane). Here we describe a fundamental relationship between these processes, suggesting that a general apparatus is used to bring about fusion, and that specificity is established by the pairing of compartment-specific proteins in the transport vesicle and target membrane, respectively.

The *N*-ethylmaleimide-sensitive fusion protein is a soluble tetramer of 76K subunits which was purified<sup>2,3</sup> on the basis of its ability to restore intercisternal Golgi transport in a cell-free

system<sup>4,5</sup>. It is required for transport vesicle fusion, as vesicles accumulate at the acceptor membrane in its absence<sup>6,7</sup>. In yeast, NSF is encoded by the *SEC18* gene<sup>8</sup>, originally shown to be necessary for transport from the endoplasmic reticulum (ER) to the Golgi<sup>9</sup>. The *SEC18* protein can replace NSF in a mammalian system for cell-free Golgi transport<sup>8</sup> and is required at every discernible step of the secretory pathway *in vivo*<sup>10</sup>. NSF thus appears to participate in a variety of intracellular fusion processes.

NSF requires additional cytoplasmic factors to attach to Golgi membranes<sup>11</sup>. Three species of monomeric soluble NSF attachment proteins, termed  $\alpha$ -,  $\beta$ - and  $\gamma$ -SNAP ( $M_s$  of 35, 36 and 39K, respectively), have been purified from brain<sup>12,13</sup> and SNAP activity is necessary for vesicle fusion *in vitro*<sup>13</sup>.  $\alpha$ -SNAP (but not  $\beta$ - or  $\gamma$ -SNAP) can restore animal cell Golgi transport activity to cytosol prepared from *sec17* mutant yeast, implying<sup>13</sup> that the *SEC17* gene encodes  $\alpha$ -SNAP in yeast. *SEC17* is now known to be functionally equivalent to  $\alpha$ -SNAP<sup>14</sup>, and the two proteins are clearly related<sup>15</sup>. In the absence of functional *SEC17* (or *SEC18*), ER to Golgi transport stops in yeast, and transport



vesicles accumulate<sup>16</sup>. SNAPs, like NSF, thus appear to be components of a general intracellular membrane fusion apparatus common to all eukaryotic cells.

SNAPs bind to distinct sites in membranes which up until now have only been operationally defined, and NSF will only interact with SNAPs that are already attached to these sites<sup>17</sup>.  $\alpha$ -SNAP and  $\beta$ -SNAP compete for binding to the same receptor site with low nM affinity.  $\gamma$ -SNAP binds to a noncompetitive site in the same complex<sup>17</sup>, and, although not essential for NSF binding, increases the complexes' affinity for NSF<sup>18</sup>. Crosslinking studies suggest that the SNAP receptor contains an  $\alpha$ -SNAP-binding subunit of 30–40K (ref. 17). When membrane-bound NSF–SNAP–SNAP receptor complexes are solubilized with detergent, they sediment as a distinct multisubunit particle at 20S (ref. 18), which may form the core of a generalized apparatus catalysing bilayer fusion at its point of assembly<sup>6,13</sup>. The 20S particles are also formed when detergent extracts of Golgi membranes containing SNAP receptors are mixed with SNAPs and NSF. When NSF and SNAPs are added in excess, all SNAP receptor activity in the membrane extract is incorporated into 20S particles<sup>18</sup>.

NSF is an ATPase containing two ATP-binding sites in separate domains<sup>19</sup>, and the binding and hydrolysis of ATP are critical in determining the stability of NSF<sup>3</sup> and its attachment to membranes<sup>2,18</sup>. Stable 20S particles can be formed in the presence of either Mg-ATP- $\gamma$ S (a non-hydrolysable analogue of ATP) or ATP without magnesium. In the presence of Mg-ATP, however, particles rapidly dissociate even at 0°C, liberating NSF in a process that requires ATP hydrolysis<sup>18</sup>. This disassembly may be an intrinsic step in the fusion mechanism<sup>18</sup>.

### Purification of SNAP receptors

To purify SNAP receptors, we used the specificity inherent in the assembly and disassembly of 20S particles as the basis of an affinity purification technique (schematized in Fig. 1): 20S particles were formed and attached to a solid matrix, allowing purified SNAP receptor to be released by particle disassembly when ATP was subsequently hydrolysed.

A recombinant form of NSF, epitope-tagged with a Myc peptide<sup>20</sup> (EQKLISEEDL) at its carboxy terminus, was expressed in *E. coli* and shown to be functional<sup>21</sup>. The 20S particles were formed in solution by mixing a Triton X-100 extract of membranes with pure, recombinant  $\alpha$ - and  $\gamma$ -SNAPs expressed in *E. coli* and NSF-Myc at 0°C (ref. 15) in the presence of ATP- $\gamma$ S and EDTA (to chelate any magnesium). A monoclonal anti-Myc IgG (9E10; ref. 22) attached to protein G beads was then added to bind the 20S particles through their NSF-Myc subunits. The anti-Myc IgG does not inhibit NSF-Myc function in cell-free transport assays (our unpublished results).

The beads were formed into a column and washed extensively before a first (nonspecific) elution with Mg-ATP- $\gamma$ S to elute any proteins attached in a Mg<sup>2+</sup>-sensitive fashion, or by Mg-ATP- $\gamma$ S *per se*. A second (specific) elution was then done with Mg-ATP (replacing Mg-ATP- $\gamma$ S). Only proteins released as

the immediate consequence of ATP hydrolysis on the column will be recovered in this specific eluate. This scheme is based on that described previously<sup>18</sup> but includes a number of key modifications detailed in the figure legends. As NSF remains on the beads, the specific eluate should consist of a fraction of the added SNAPs, together with additional polypeptides representing SNAP receptors, ideally in stoichiometric amounts.

A detergent extract of a crude, salt-washed total particulate fraction from the grey matter of bovine brain was used as the source of potential SNAP receptors. Membranes were washed with 1 M KCl before use to remove most of their endogenous SNAP supply<sup>12</sup>. As recombinant SNAPs were added in great excess for the binding reaction (Fig. 1), these should compete out the remaining endogenous SNAPs, preventing them from forming particles on the beads. Figure 2a shows a Coomassie blue-stained SDS-urea-high Tris-polyacrylamide gel of the specific eluate (lane 3), the nonspecific eluate fraction (lane 2), and the specific eluate from a control experiment omitting NSF-Myc (lane 1). Several bands (labelled A–F) appear only in the specific (Mg-ATP) eluate and depend on the presence of NSF. A set of bands at about  $M_r$  70K are present in all three lanes, but apart from these, bands A–F are substantially pure in the specific eluate fraction. Band F runs as a sharper band in a standard Laemmli gel (Fig. 2b; specific eluate, stained with silver) than in the high Tris-urea gel (Fig. 2a). Bands A and C virtually co-electrophorese with the abundant band D in a Laemmli gel (Fig. 2b; see below).

### Identification of SNAP receptors

To identify specific bands, they were excised from blots and digested with trypsin, giving peptides that were then separated by high-pressure liquid chromatography and microsequenced (see Fig. 3 for details). Bands B and D co-electrophoresed with recombinant  $\gamma$ -SNAP and  $\alpha$ -SNAP, respectively, both of which were His<sub>6</sub>-tagged<sup>15</sup> and thus slightly larger than their endogenous counterparts (not shown). These identifications were confirmed by microsequencing peptides from bands B and D (not shown).

Sequences from six major peptides containing a total of 70 residues were obtained from band A, each precisely matching the sequence of syntaxin B (p35B) from rat brain<sup>23</sup> (Fig. 3b). The sequences of these and peptides from the other specific bands are shown in association with the band from which they were obtained (Fig. 3a). No sequences or fragments attributable to any other protein were found, indicating that the bulk of material in band A is syntaxin B.

Five major peptides containing a total of 53 residues were sequenced from band C (Fig. 3a), and proved to differ at only one position from the published sequence of syntaxin A from rat brain<sup>23</sup>; the difference presumably reflects the species involved. Expected sequence differences between the 84% identical syntaxins A and B were found in the peptides from bands C and A (Fig. 3b). The syntaxins run more slowly than expected from their relative molecular masses (about 35K) in the high percentage Tris-SDS-urea-polyacrylamide gels were used,

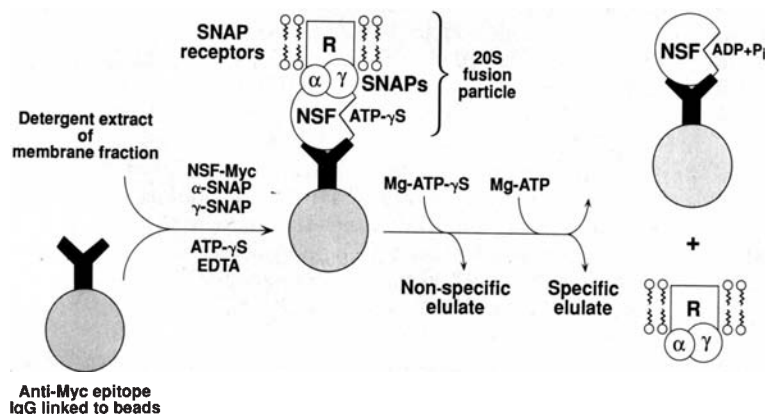


FIG. 1 Procedure used to purify SNAP receptors (SNAREs). Recombinant NSF,  $\alpha$ -SNAP, and  $\gamma$ -SNAP are assembled into 20S particles by SNAREs present in a crude detergent extract of membranes. The SNAREs are incorporated stoichiometrically into the particles, which are then bound to beads by means of NSF. For this purpose, the NSF is epitope-tagged with Myc, and an anti-Myc monoclonal IgG is linked to the beads. The beads are washed and then eluted first with Mg-ATP- $\gamma$ S (nonspecific eluate) and then with Mg-ATP (specific eluate). The bound 20S particles disassemble in the presence of Mg-ATP (but not Mg-ATP- $\gamma$ S), releasing stoichiometric amounts of SNAPs and SNAREs. NSF remains bound to the beads. Experimental details are given in Fig. 2 legend.

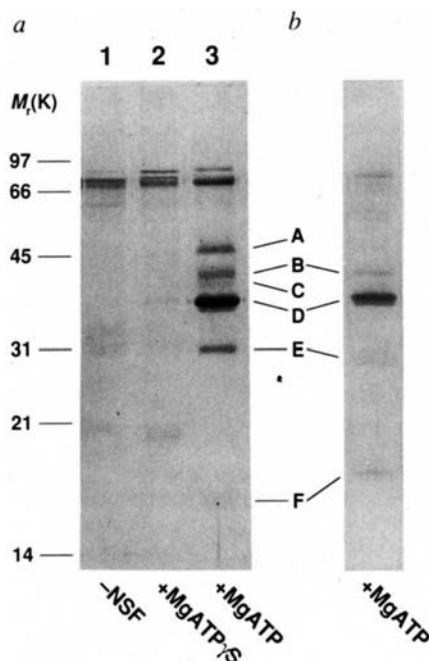


FIG. 2 Identification of proteins released from NSF after ATP hydrolysis. *a*, Polyacrylamide gel stained with Coomassie blue. Lane 1, control, Mg-ATP eluate of control binding reaction in the absence of NSF; lane 2, 'nonspecific' eluate from complete binding reaction with NSF-Myc and Mg-ATP- $\gamma$ S; lane 3, 'specific' eluate of the same column as for lane 2 following the exchange of ATP for ATP- $\gamma$ S (in the presence of EDTA) and addition of  $Mg^{2+}$  to allow ATP hydrolysis (Fig. 1). *b*, Silver-stained Laemmli gel of the specific (Mg-ATP) eluate.

**METHODS.** All manipulations were performed at 0–4 °C. Bovine brain tissue was initially stripped of meninges, and all but the grey matter removed and discarded. The resulting tissue (30g) was homogenized with 30 strokes of a Dounce homogenizer in buffer A (20 mM Tris-HCl, pH 8.0, 1 M KCl, 250 mM

although they virtually co-electrophorese with  $\alpha$ -SNAP (35K) in standard Laemmli gels (Fig. 2*b*, and data not shown).

Three major peptides (Fig. 3*a*) from band E gave sequences (Fig. 3*c*) matching a protein predicted from a mouse brain complementary DNA clone which by coincidence has been termed SNAP-25 (for synaptosome-associated protein of 25K; ref. 24). Of 45 residues obtained, 42 were identical to the published sequence of mouse SNAP-25. The discrepancies, two of which are conservative, are probably due to species differences. SNAP-25 migrates at  $M_r$  31K in the SDS-urea gel.

All five peptides (containing 43 identified residues) sequenced from band F (Fig. 3*a*) exactly matched the sequence of VAMP/synaptobrevin-2 from bovine brain<sup>25</sup> (Fig. 3*d*). VAMP-2, the major isoform of VAMP/synaptobrevin in brain, differs slightly from VAMP-1 (refs 26, 27); all the sequences obtained from band F correspond to VAMP-2. All HPLC peptide peaks from band F, other than those attributable to trypsin autolysis, were shown to be derived from VAMP/synaptobrevin. As a control, the region corresponding to band F from the blot of the gel of the nonspecific (Mg-ATP- $\gamma$ S) eluate was subject to tryptic digestion and HPLC analysis in parallel with band F from the Mg-ATP eluate in two separate experiments. None of the VAMP/synaptobrevin-2 peptide peaks from band F were present in the HPLC profile of the control; only the autolytic tryptic peptides were detected.

Two peptides from the digest of band F failed to give any sequence, suggesting that they had blocked N termini. Mass spectroscopy indicated that these components had  $m/z$  values of 2,680.90 and 2,838.36 respectively, differing by 157.46, or approximately the molecular mass of Arg (156.19 average isotopic mass). The tryptic cleavage site closest to the N terminus in bovine VAMP/synaptobrevin-2 consists of an Arg-Arg sequence (at positions 30–31), so the expected partial cleavage

sucrose, 2 mM  $MgCl_2$ , 1 mM DTT, 1 mM phenylmethylsulphonyl fluoride (PMSF). A total particulate fraction was isolated by centrifugation in a Ti45 rotor (Beckman) for 60 min at 33,000 r.p.m. The pellet was then resuspended in buffer A by Dounce homogenization and the membranes collected by centrifugation. The resulting pellet was washed once in buffer B (10 mM HEPES/KOH, pH 7.8, 100 mM KCl, 2 mM  $MgCl_2$ , 1 mM DTT) and then resuspended in 100 ml buffer B. Triton X-100 was added slowly with mixing to a final concentration of 4% (v/v), and the suspension incubated on ice with frequent mixing. After 45 min, the suspension was clarified by centrifugation (Ti-45 rotor) for 60 min at 33,000 r.p.m. and the supernatant dialysed overnight against 100 vol 25 mM Tris-HCl, pH 7.8, 50 mM KCl, 1 mM DTT, 1% (v/v) Triton X-100. After dialysis, the material was clarified by centrifugation (Ti-45 rotor for 60 min at 33,000 r.p.m.), aliquoted, and stored at –80 °C. Protein concentration was measured using the BCA reaction (Pierce) with ovalbumin as standard. This bovine brain extract (2mg protein) was preincubated in the presence of His $_6$ - $\alpha$ -SNAP (12  $\mu$ g protein), His $_6$ - $\gamma$ -SNAP (4  $\mu$ g protein) and NSF-Myc (12  $\mu$ g protein) in buffer C (25 mM HEPES/KOH, pH 7.0, 0.75% (w/v) Triton X-100, 75 mM KCl, 1 mM DTT, 2 mM EDTA, 0.5 mM ATP (for lane 1) or 0.5 mM ATP- $\gamma$ S (for lanes 2 and 3), 1% (w/v) polyethylene-glycol (PEG) 4,000, 0.4 mM PMSF) for 30 min at 4 °C in a final volume of 2 ml. Mouse anti-Myc monoclonal antibody (200  $\mu$ g protein, termed 9E10; ref. 22) covalently coupled to protein G-Sepharose 4 fast-flow (Pharmacia) was added and the incubation continued for 2 h with constant agitation. (The anti-Myc IgG was covalently coupled to the protein G-Sepharose using 20 mM dimethylsuberimide as described<sup>47</sup>.) The beads were packed into a column, washed with 10 vol buffer D (20 mM HEPES/KOH, pH 7.0, 0.5% (w/v) Triton X-100, 100 mM KCl, 1 mM DTT, 2 mM EDTA, 0.5 mM ATP (lane 1) or 0.5 mM ATP- $\gamma$ S (lanes 2 and 3), 0.4 mM PMSF) and the proteins eluted with 6 column-volumes of buffer D containing 8 mM  $MgCl_2$  (resulting in a concentration of free  $Mg^{2+}$  of 4 mM). The column containing ATP- $\gamma$ S (lanes 2 and 3) was washed with an additional 2 column-volumes of buffer D containing 4 mM EDTA to complex the free  $Mg^{2+}$ , 3 column-volumes of buffer D containing ATP to exchange for bound ATP- $\gamma$ S, and then eluted with 6 column-volumes of buffer D containing 8 mM  $MgCl_2$  to allow ATP hydrolysis. The appropriate fractions containing the eluate were pooled, precipitated with trichloroacetic acid, boiled in sample buffer<sup>48</sup> and analysed by electrophoresis on Tris-urea/SDS-polyacrylamide (18% acrylamide, 6M urea, 750 mM Tris-HCl, pH 8.85, 50 mM NaCl, 0.1% SDS<sup>49</sup>) and stained with Coomassie blue R-250. For sequencing, the reaction was scaled up 25-fold.

would yield two N-terminal-derived peptides. If it is assumed that the initiator Met is removed, and that the newly generated N-terminal Ser is acetylated, adding 42.04 to the  $M_r$ , the two masses correspond almost perfectly with the predicted  $M_s$  [ $MH^+$ ] of the bovine VAMP/synaptobrevin-2 peptides spanning residues 2–30 and 2–31 (2,681.90 and 2,838.09, respectively). The N terminus of VAMP/synaptobrevin-2 thus appears to be processed and acetylated.

The material in the 70K region of all three eluates (specific, nonspecific and control) (Fig. 2) was microsequenced (from the specific eluate) in an attempt to determine if any synaptotagmin/p65 was present, as this protein can be immunoprecipitated together with syntaxins<sup>23</sup>. No evidence for its presence was found, although it cannot be excluded. Of three peptides sequenced, two were from Hsp70, an ATP-binding protein<sup>28</sup>, and one was from a protein not found in the Genbank database.

Scanning and integrating the Coomassie blue stain in the gel shown in Fig. 2*a* gives the following ratios when staining is corrected for molecular weight, expressed as a fraction of  $\alpha$ -SNAP: syntaxin B, 0.18; SNAP-25, 0.15; VAMP/synaptobrevin-2, 0.23. The syntaxin A band (band C) is weak compared with syntaxin B and could not readily be quantitated. The sum of all of the SNAP receptor species is 0.56 mol per mol  $\alpha$ -SNAP, and so is approximately stoichiometric, given that proteins vary in staining by Coomassie blue, and the stoichiometry of SNAP and receptor in complexes does not need to be 1:1. The relative amounts of the different SNAP receptors in such a co-purified mixture may not reflect their relative abundance in the starting material, as their affinities for SNAPs may differ. The ratio of  $\gamma$ -SNAP to  $\alpha$ -SNAP was ~0.2.

The role of syntaxins, SNAP-25, and synaptobrevin as SNAP receptors is supported by (1) their absence when NSF was omitted from the purification; (2) the need for ATP hydrolysis,





20S particles from SNAPs and NSF, and their release from NSF upon ATP hydrolysis (studies using Golgi membranes as source material will be described elsewhere). The simplest explanation for the fact that multiple polypeptides are obtained is that each is a distinct SNAP receptor. In particular, there is no evidence that any of these polypeptides associate, for example, as heterodimers<sup>54</sup>. All of them must thus have either the capacity to assemble a 20S particle, or a very high affinity for one that has been assembled. As our method is biased to isolating receptors of the highest abundance and affinity, we have probably isolated only some of the receptor types that are present.

By exploiting the same enzyme cycle used in the vesicle fusion process to offer two layers of biological specificity, we have purified SNAP receptors on the basis of their function, giving a largely pure preparation of four proteins from the crudest possible extract of brain, all of which originate from synapses. This striking selection for synaptic components probably reflects the degree to which the brain is specialized for synaptic vesicle fusion. The fact that each SNAP receptor purified corresponds to a previously cloned and sequenced gene no doubt reflects the exhaustive structural characterization of synaptic components by molecular biologists who have focused on the synapse because of its central importance in neuronal development and function<sup>29,30,54</sup>. Despite this effort, however, the roles of these proteins are unknown because of the lack of functional assays.

Of the SNAP receptors we have isolated, SNAP-25 is found in the presynaptic terminals of neurons<sup>24</sup> but has not been more precisely localized. Although its sequence suggests it is hydrophilic, it behaves as an integral membrane protein<sup>24</sup> and is palmitoylated at one or more of the four cysteine residues between positions 85 and 92 (ref. 55). Fatty acyl-CoA is required for NSF-dependent fusion<sup>56</sup>, and SNAP-25 or related proteins may serve as the relevant acyl acceptors. Multiple acylations of such proteins could create a hydrophobic surface as a trigger for fusion.

VAMP/synaptobrevin is inserted into synaptic vesicles by a single hydrophobic C-terminal segment, and the remainder of

the protein is cytoplasmic. Both tetanus and botulinum B toxins, potent inhibitors of neurotransmitter release, are proteases specific for VAMP/synaptobrevin-2 (ref. 31), suggesting that this SNAP receptor is essential for synaptic vesicle fusion *in vivo*. Syntaxin<sup>23</sup> has the same topography as VAMP/synaptobrevin, but is concentrated in 'active zones' of the presynaptic membrane<sup>23</sup> at which a subpopulation of synaptic vesicles is docked<sup>32</sup> to allow fusion within 200  $\mu$ s of the calcium influx induced by an action potential<sup>33</sup>. Components of the fusion machinery may have to be largely preassembled at these sites to allow such a rapid response.

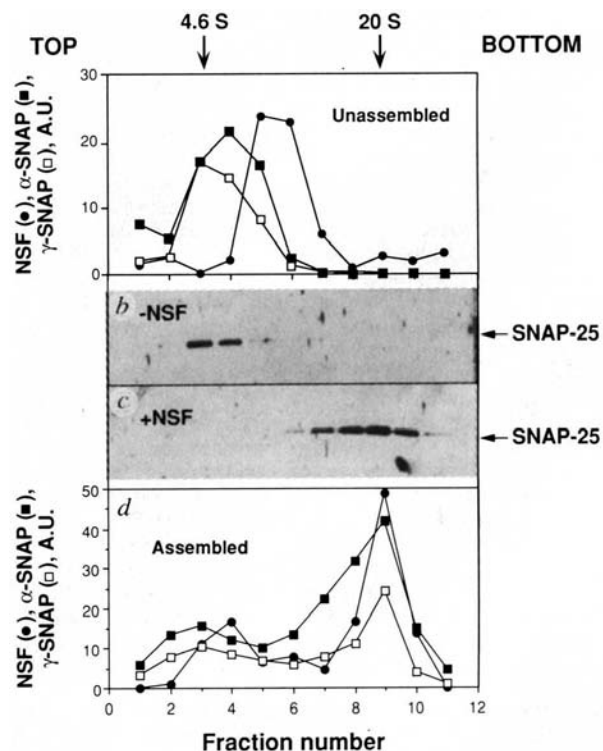
In this light, the fact that 20S particles can contain VAMP/synaptobrevin (from the synaptic vesicle) and syntaxin (from the plasma membrane) is of interest, because these complexes might form part of the fusion apparatus docking the vesicle to its target in the active zone. It is not yet possible to say whether syntaxin and VAMP/synaptobrevin exist together in a single 20S particle (Fig. 5a), or whether separately assembled particles join together to form an attachment site (Fig. 5b), perhaps with the aid of additional proteins.

The fact that molecules previously implicated in regulated exocytosis at the synapse are SNAP receptors implies that the same NSF- and SNAP-dependent machinery necessary for many constitutive fusion events also underlies triggered release of neurotransmitters at synapses. By extension, we would expect that NSF and SNAP are similarly employed in other forms of regulated exocytosis<sup>34</sup>.

How can constitutive fusion machinery also be used in triggered exocytosis? An inhibitory component(s) must apparently prevent the 20S fusion particle from engaging, or from completely assembling in the first place. Such a fusion clamp could either cover an active site in NSF, SNAP or the SNAP receptors, or could prevent any of the additional cytoplasmic subunits of the fusion machinery<sup>35</sup> from binding. In the case of the synapse, a calcium trigger would remove the clamp. A strong candidate for such a clamp is the synaptic vesicle membrane protein synaptotagmin<sup>36,37</sup>. Synaptotagmin co-immunoprecipitates with syntaxin<sup>23</sup>, and the recombinant pro-

FIG. 4 Incorporation of SNAP-25 and other components into 20S fusion particles. *a, b*, Conditions in which 20S particles do not form (for example, controls); *c, d*, conditions in which 20S particles assemble. Components of the 20S fusion particle were incubated and then fractionated by glycerol gradient centrifugation, and the fractions analysed by SDS-PAGE to determine the location of  $\alpha$ -SNAP and  $\gamma$ -SNAP, and to determine the locations of NSF and SNAP-25. Bovine serum albumin (4.6S) and  $\alpha_2$ -macroglobulin (20S) were used as standards.

**METHODS.** Reaction conditions were arranged so that the protein of interest would be maximally incorporated into 20S particles; thus, results in *a* and *d* are composites of several experiments. In the case in which the SNAPs were examined, *in vitro*-translated <sup>35</sup>S-labelled  $\alpha$ -SNAP (■) and  $\gamma$ -SNAP (□) (ref. 17) ( $\sim 2.1 \times 10^6$  c.p.m. for  $\alpha$ - and  $1.4 \times 10^5$  c.p.m. for  $\gamma$ -SNAP) were incubated on ice with bovine brain extract (1.2 mg protein), in the absence (*a*) or presence (*d*) of His<sub>6</sub>-NSF (100  $\mu$ g protein) in 20 mM HEPES-KOH, pH 7.4, 100 mM KCl, 2 mM DTT, 2 mM EDTA, 0.5 mM ATP, 0.5% (v/v) Triton X-100 in a final volume of 0.5 ml. After 15 min, reactions were loaded onto a 10–35% (w/v) glycerol gradient<sup>18</sup> and centrifuged in an SW41 rotor (Beckman) for 18 h at 40,000 r.p.m. Gradients were fractionated from the bottom at 1 ml per min, and an aliquot of each fraction analysed by SDS-PAGE and autofluorography. Recovery of both radiolabelled SNAPs exceeded 90%. Autofluorographs were scanned with a ScanJet Plus (Hewlett Packard) and the images integrated using Scan Analysis software (BioSoft, Cambridge). To examine NSF (●), His<sub>6</sub>-NSF (ref. 15) (10  $\mu$ g protein) was incubated with each of the His<sub>6</sub>-SNAPs (20  $\mu$ g each protein), either in the absence (*a*) or presence (*d*) of bovine brain extract (1.5 mg protein). Samples were incubated and fractionated as described. Aliquots were analysed by western blotting using an anti-His<sub>6</sub> antibody and blots were scanned as described<sup>17</sup>. To determine the extent to which SNAP-25 is incorporated into the 20S particle, bovine brain extract was added in limiting amounts (300  $\mu$ g protein) and incubated either in the absence (*b*) or presence (*c*) of both His<sub>6</sub>-SNAPs and His<sub>6</sub>-NSF (30  $\mu$ g each protein). Samples were processed and aliquots of each fraction analysed by western blotting using an antibody generated against the C-terminal peptide of SNAP-25. For this purpose, a peptide of



the sequence used in ref. 24 was synthesized but with an additional cysteine at the N terminus, and affinity-purified antibodies generated. Blots were visualized using the ECL System (Amersham). AU, arbitrary unit.



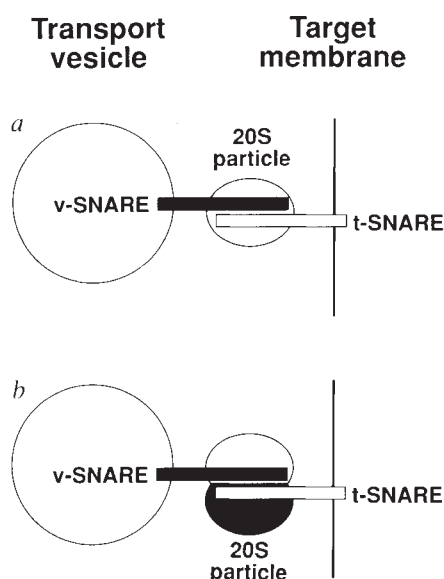


FIG. 5 Models to explain vesicle targeting based on the finding that SNAREs isolated in 20S fusion particles can originate from either the transport vesicle (v-SNAREs) or from the target membrane (t-SNAREs). A 20S particle (containing NSF and SNAPs) that simultaneously binds a v-SNARE and a t-SNARE (a) would attach a vesicle to its target. Alternatively (b), 20S particles, each capable of binding only one SNARE at a time, could interact to attach vesicle to target, a process that perhaps requires other proteins to assemble together.

teins can bind each other<sup>23</sup>. It also undergoes a calcium-dependent conformational change<sup>38</sup>. Although synaptotagmin is not one of the major components of the 20S particles characterized here, this is not surprising, as the interaction between syntaxin and synaptotagmin is disrupted by 0.5 M salt<sup>23</sup>, and we used a 1 M salt wash.

Rab3A, a Ras-related GTP-binding protein, is another candidate for a clamp, as this protein dissociates from synaptic vesicles upon stimulation of fusion<sup>39</sup>. Fusion clamps may respond to a variety of second messengers such as cyclic AMP, activated G proteins, protein phosphorylation, and so on, in addition to calcium, depending on cell type and physiological context.

### SNAREs may encode specificity

Both syntaxin and VAMP/synaptobrevin have homologues in yeast<sup>40,42</sup>. The yeast protein SED5 is related to syntaxin<sup>40</sup>, and its absence blocks ER-to-Golgi transport, leading to the accumulation of transport vesicles; it appears to reside in the Golgi<sup>40</sup>. Another homologue, PEP12, is found primarily on the vacuole membrane, and is required for Golgi-to-vacuole transport<sup>41</sup>. The yeast *SEC22*, *BET1* and *BOS1* gene products are related to VAMP/synaptobrevin, and are required for fusion with the Golgi *in vivo*. *BOS1* at least is a component of ER-derived transport vesicles<sup>16,42-45</sup>. Other recent examples are discussed in ref. 46. SNAP-25 is distantly related to SED5, PEP12, and the syntaxins.

The synaptic SNAP receptors we have identified are thus members of a family of proteins which appear to be distributed in a compartmentally specific fashion, with one set attached to the transport vesicles (v-SNAREs) and another set attached to target membranes (t-SNAREs). This suggests a model in which each transport vesicle contains one or more members of the v-SNARE superfamily obtained on budding from a corresponding donor compartment, and every target compartment in a cell contains one or more members of the t-SNARE superfamily. Specificity in membrane transactions would be assured by the unique and non-overlapping distribution of v-SNAREs and t-SNAREs among the different vesicles and target compartments. In the simplest view, that is, if there were no other source of

specificity, only when complementary v-SNARE and t-SNARE pairs engage would a productive fusion event be initiated. Thus, a v-SNARE from the ER (possibly SEC22) could engage a t-SNARE from the Golgi (possibly SED5), but not one from the lysosome (possibly PEP12). For the proposed mechanism to work (again, assuming no other source of specificity), non-complementary SNAREs would have to be excluded from the same 20S particle (Fig. 5a) or multi-particle assemblies (Fig. 5b), although any one SNARE might be able to form a 20S particle by itself while waiting for a partner to arrive. If this is correct, the 20S fusion particle assembly reaction might be used as a cell-free 'read-out' system to test whether candidate proteins are SNAREs (as shown here for SNAP-25), and to establish which SNAREs, if any, form specific cognate pairs. Localization of the SNARE proteins *in situ* would then establish which are v-SNAREs and which are t-SNAREs and which compartments are involved in each pairing.

A related mechanistic question is whether the SNAPs (and NSF) are associated primarily with a v-SNARE or with a t-SNARE, or whether an essentially symmetrical structure is formed at the junction of vesicle and target. NSF and SNAPs can associate with Golgi membranes to form 20S particles in the absence of transport vesicles<sup>18</sup>, implying that 20S particles can be formed with only one SNARE partner. As native NSF is a tetramer of subunits with two ATPase domains each, it is easy to imagine several ways in which a 20S particle containing one NSF could form a symmetrical vesicle-target junction by simultaneously binding a v-SNARE and a t-SNARE (Fig. 5a).

Although many important details need to be established, our findings imply a general role for NSF and SNAP in regulated and constitutive intracellular membrane fusion processes, and in synaptic transmission in particular. The SNAREs we have identified appear to be members of compartment-specific membrane protein mutigene families which may form attachment sites between specific vesicles and their correct target membranes together with NSF and SNAPs. □

Received 25 February; accepted 9 March 1993.

- Palade, G. E. *Science* **189**, 347-358 (1975).
- Glick, B. S. & Rothman, J. E. *Nature* **326**, 309-312 (1987).
- Block, M. R., Glick, B. S., Wilcox, C. A., Wieland, F. T. & Rothman, J. E. *Proc. natn. Acad. Sci. U.S.A.* **85**, 7852-7856 (1988).
- Fries, E. & Rothman, J. E. *Proc. natn. Acad. Sci. U.S.A.* **77**, 3870-3874 (1980).
- Balch, W. E., Dunphy, D. W., Braell, W. A. & Rothman, J. E. *Cell* **39**, 405-416 (1984).
- Malhotra, V., Orci, L., Glick, B. S., Block, M. R. & Rothman, J. E. *Cell* **54**, 221-227 (1988).
- Orci, L., Malhotra, V., Amherdt, M., Serafini, T. & Rothman, J. E. *Cell* **56**, 357-368 (1989).
- Wilson, D. W. *et al. Nature* **339**, 355-359 (1989).
- Novick, P., Ferro, S. & Schekman, R. *Cell* **25**, 461-469 (1981).
- Graham, T. R. & Emr, S. D. *J. Cell Biol.* **114**, 207-218 (1991).
- Weidman, P. J., Melancon, P., Block, M. R. & Rothman, J. E. *J. Cell Biol.* **108**, 1589-1596 (1989).
- Clary, D. O. & Rothman, J. E. *J. Biol. Chem.* **265**, 10109-10117 (1990).
- Clary, D. O., Griff, I. C. & Rothman, J. E. *Cell* **61**, 709-721 (1990).
- Griff, I. C., Schekman, R., Rothman, J. E. & Kaiser, C. A. *J. Biol. Chem.* **267**, 12106-12115 (1992).
- Whiteheart, S. W. *et al. Nature* **362**, 353-355 (1993).
- Kaiser, C. A. & Schekman, R. *Cell* **61**, 723-733 (1990).
- Whiteheart, S. W., Brunner, M., Wilson, D. W., Wiedmann, M. & Rothman, J. E. *J. Biol. Chem.* **267**, 12239-12243 (1992).
- Wilson, D. W., Whiteheart, S. W., Wiedmann, M., Brunner, M. & Rothman, J. E. *J. Cell Biol.* **117**, 531-538 (1992).
- Tagaya, M., Wilson, D. W., Brunner, M., Arango, N. & Rothman, J. E. *J. Biol. Chem.* **269**, 2662-2666 (1993).
- Munro, S. & Pelham, H. R. B. *Cell* **48**, 899-907 (1987).
- Wilson, D. W. & Rothman, J. E. *Meth. Enzym.* **219**, 309-318 (1992).
- Evan, G. I., Lewis, G. K., Ramsey, G. & Bishop, M. J. *Molec. cell. Biol.* **5**, 3610-3616 (1985).
- Bennett, M. K., Calakos, N. & Scheller, R. H. *Science* **257**, 255-259 (1992).
- Oyler, G. A. *et al. J. Cell Biol.* **109**, 3030-3052 (1989).
- Sudhof, T. C. *et al. Neuron* **2**, 1475-1481 (1989).
- Elferink, L. A., Trimble, W. S. & Scheller, R. H. *J. Biol. Chem.* **264**, 11061-11064 (1989).
- Archer, B. T., Ozelik, T., Jahn, R., Franke, U. & Sudhof, T. C. *J. Biol. Chem.* **265**, 17267-17273 (1990).
- Chappell, T. G. *Cell* **45**, 3-13 (1986).
- Greengard, P., Valtorta, F., Czernik, A. J. & Benfenati, F. *Science* **259**, 780-785 (1993).
- Trimble, W. S., Linal, M. & Scheller, R. H. *Rev. Neurosci.* **14**, 93-122 (1991).
- Chapman, G. *et al. Nature* **359**, 832-835 (1992).
- Landis, D. M. D., Hall, A. K., Weinstein, L. A. & Reese, T. S. *Neuron* **1**, 201-209 (1988).
- Linas, R., Steinberg, I. Z. & Walton, K. *Biophys. J.* **33**, 323-352 (1981).
- Kelley, R. B. *Science* **230**, 25-32 (1985).
- Waters, R. G., Clary, D. O. & Rothman, J. E. *J. Cell Biol.* **118**, 1015-1026 (1992).
- Perin, M. S., Fried, V. A., Mignery, G. A., Jahn, R. & Sudhof, T. C. *Nature* **34**, 260-263 (1990).
- Wendland, B., Miller, K. G., Schilling, J. & Scheller, R. H. *Neuron* **6**, 993-1007 (1991).
- Brose, N., Petrenko, A. G., Sudhof, T. C. & Jahn, R. *Science* **256**, 1021-1025 (1992).
- Fischer von Mollard, G., Sudhof, T. C. & Jahn, R. *Nature* **359**, 79-81 (1991).
- Hardwick, K. G. & Pelham, H. R. B. *J. Cell Biol.* **119**, 513-521 (1992).
- Preston, R. A. *et al. Molec. cell. Biol.* **11**, 5801-5812 (1991).

42. Dascher, C., Ossig, R., Gallwitz, D. & Schmitt, H. D. *Molec. cell. Biol.* **11**, 872–885 (1991).
43. Newman, A. P., Groesch, M. E. & Ferro-Novick, S. *EMBO J.* **11**, 3609–3617 (1992).
44. Newman, A. P. *et al.* *Molec. cell. Biol.* **12**, 3663–3664 (1992).
45. Shim, J., Newman, A. P. & Ferro-Novick, S. *J. Cell Biol.* **113**, 55–64 (1991).
46. Bennett, M. K. & Scheller, R. H. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
47. Harlow, E. & Lane, O. *Antibodies: A Laboratory Manual* (Cold Spring Harbor Laboratory Press, New York, 1988).
48. Laemmli, U. K. *Nature* **227**, 680–685 (1970).
49. Schlenstedt, G., Gudmundsson, G. H., Boman, H. G. & Zimmermann, R. *J. biol. Chem.* **265**, 13960–13968 (1990).
50. Tempst, P., Link, A. J., Riviere, L. R., Fleming, M. & Elicone, C. *Electrophoresis* **11**, 537–553 (1990).
51. Tempst, P. & Riviere, L. *Analyt. Biochem.* **183**, 290–300 (1989).
52. Erdjument-Bromage, H., Geromanos, S., Chodera, A. & Tempst, P. *Techniques in Protein Chemistry IV* (Academic, San Diego, in the press).
53. Beavis, R. C. & Chait, B. T. *Rapid Commun. Mass Spect.* **3**, 432–435 (1989).
54. Sudhof, T. C. & Jahn, R. *Neuron* **6**, 665–677 (1991).
55. Hess, D. T., Slater, T. M., Wilson, M. C. & Skene, J. H. P. *J. Neurosci.* **12**, 4634–4641 (1992).
56. Pfanner, N., Glick, B. S., Arden, S. R. & Rothman, J. E. *J. Cell Biol.* **110**, 955–961 (1990).

ACKNOWLEDGEMENTS. We thank M. Hum, L. Lacomis and M. Lui for help with protein digestions, HPLC separations and peptide sequencing; C. Elicone for custom assembly of the microbore HPLC system; M. Wiedmann for discussion; and W. Patton for gel scans and analysis. This work was supported by the Mathers Charitable Foundation, and by an NIH grant (to J.E.R.), a Fellowship of the Deutsche Forschungsgemeinschaft (to T.S.), a Fellowship from the Jane Coffin Childs Memorial Fund for Medical Research (to S.W.W.), and by an EMBO postdoctoral fellowship (to M.B.). The Sloan-Kettering microchemistry core facility is supported, in part, by an NCI core grant.

## LETTERS TO NATURE

## A young source of optical emission from distant radio galaxies

F. Hammer<sup>\*†</sup>, O. Le Fèvre<sup>\*†</sup> & M. C. Angonin<sup>\*</sup>

<sup>\*</sup>Département d'Astrophysique Extragalactique et Cosmologie, Observatoire de Paris-Meudon, 92195 Meudon Cedex, France

<sup>†</sup>Canada-France-Hawaii Telescope Corporation, PO Box 1597, Kamuela, Hawaii 96743, USA

**DISTANT radio galaxies provide valuable insights into the properties of the young Universe—they are the only known extended optical sources at high redshift and might represent an early stage in the formation and evolution of galaxies in general. This extended optical emission often has very complex morphologies, but the origin of the light is still unclear. Here we report spectroscopic observations for several distant radio galaxies ( $0.75 \leq z \leq 1.1$ ) in which the rest-frame spectra exhibit featureless continua between 2,500 Å and 5,000 Å. We see no evidence for the break in the spectrum at 4,000 Å expected for an old stellar population<sup>1–3</sup>, and suggest that young stars or scattered emissions from the active nuclei are responsible for most of the observed light. In either case, this implies that the source of the optical emission is comparable in age to the associated radio source, namely  $10^7$  years or less.**

One of the keys to understanding the nature of distant radio galaxies is the study of their optical continuum, and multiple broad-band photometry has been extensively used to examine their spectral energy distribution, often in conjunction with stellar population synthesis models. Contamination from various sources, however, can hide the true properties, and obviously affects the interpretation of the data. For example, distant radio galaxies generally have very strong emission lines, which can contribute significantly to broad-band photometric measurements both at visible and near-infrared wavelengths<sup>4,5</sup> (equivalent widths larger than several hundred angstroms); foreground galaxies<sup>6</sup> and even galactic stars<sup>7</sup> can project on the lines of sight. This is especially true when the broad-band photometry is performed on high redshift radio galaxies whose physical origin is still unclear and for which the multiple broad-band photometric data are at least as well fitted by simple power laws. To circumvent the problems inherent in such photometry<sup>6–8</sup>, our goal here is to test the stellar contribution to the optical emission of distant radio galaxies from deep spectroscopic data.

The most convincing (stellar) model was presented by Lilly<sup>1</sup> and developed by Rigler *et al.*<sup>2</sup>. It assumes that the red continuum emission is due to a symmetric giant elliptical galaxy which becomes dominant at observed infrared wavelengths, while the ultraviolet emission consists of a flat component responsible for the alignment with the radio axis. Rigler *et al.*<sup>2</sup> have called this interpretation the 'old star hypothesis', and named another stellar scheme developed by Chambers and

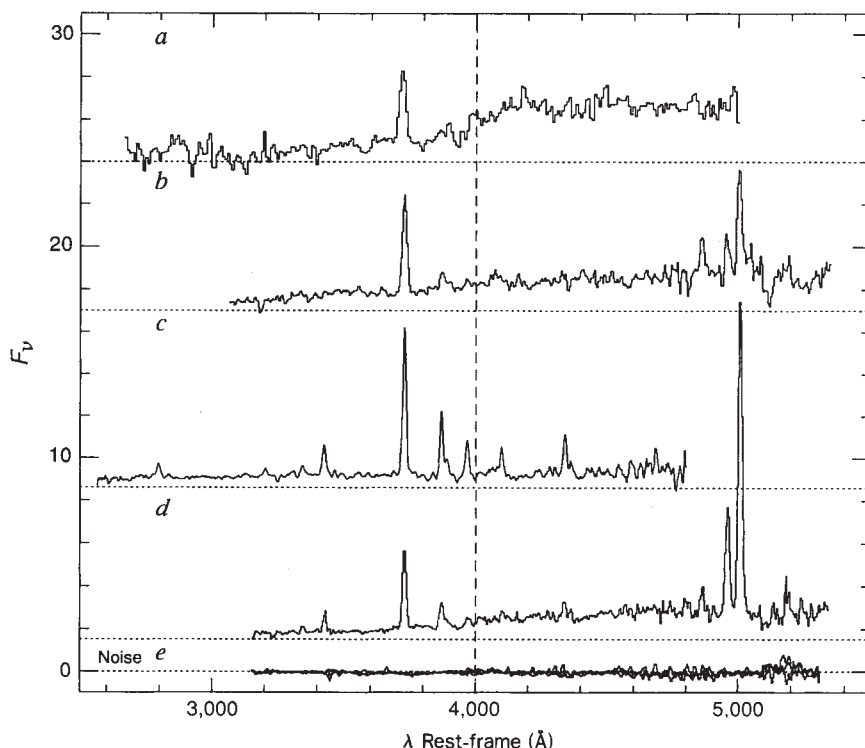
Charlot<sup>3</sup> the 'range of ages model'. The latter assumes a significant contribution of red supergiant stars (asymptotic giant branch) to the initial mass function, and that the optical colours of distant radio galaxies are an indication of their ages. Both models are apparently compatible with the small scatter of the K-band-magnitude-redshift ( $K-z$ ) diagram, and seem adequately to fit the multiple broad-band photometric properties of distant radio galaxies. With the blue and red populations contributing roughly equally to the flux at 4,000 Å (ref. 2) in the 'old star' hypothesis, the continuum should exhibit a 4,000 Å break (from the old stars ( $>1$  Gyr)) with a slope flattening due to the contribution of the blue population. Similarly, the 'range of ages' model predicts a significant 4,000 Å break for all sources, and especially for the reddest ones. Another model, put forward by Bithell and Rees<sup>9</sup>, assumes a single and extremely young stellar component for distant radio galaxies and expects spectral signatures from young stars between 3,500 and 4,500 Å.

We have obtained deep red spectroscopy of a sample of distant radio galaxies with redshifts larger than 0.7 (beyond which the alignment effect becomes significant), and smaller than 1.1 (beyond which features above 4,700 Å enter the infrared window). One can successfully extract spectra of galaxies with red-magnitude ( $R$ ) = 21.5 with a signal to noise ratio ( $S/N$ )  $\geq 10$  around 9,000 Å on 4-m class telescopes in 3 hours, with an excellent uniformity of the sky correction (that is, an average pixel-to-pixel deviation less than 10% of the corrected  $R = 21.5$  continuum, or less than 0.1% of the sky background (O.L.F., manuscript in preparation)). The ability to detect features in the continuum of faint galaxies beyond  $z = 1$  was demonstrated from our spectroscopy of the quasar 3CR245 companion<sup>10</sup>. This is a radio-quiet galaxy at  $z = 1.013$ , which has optical magnitudes more than 1 mag less than distant 3CR galaxies (visible magnitude  $V = 22.4$ ), and for which the 4,000 Å break and the Ca I, H and K bands are remarkably well detected in only two integrations of 45 min each ( $S/N > 10$ ). The galaxy morphology and the amplitude of the break ( $D(4,000 \text{ Å}) = 2$ ; see Bruzual<sup>11</sup> for the definition) are characteristic of an elliptical galaxy, as is the detection of relatively strong absorption lines (Ca II, G band). For 3CR galaxies which have brighter magnitudes at optical wavelengths, one would therefore expect to identify spectral absorption features in the continuum in only a few hours of integration under similar observing conditions.

We observed 10 distant 3CR galaxies with redshifts ranging from 0.75 to 1.1. Their optical emission is aligned within 30° of the radio axis, and their infrared-visual colours range from  $R-K = 2.5$  to 4, typical of distant radio galaxies. Some of our objects are red enough to be classified by Chambers and Charlot<sup>3</sup> as older galaxies ( $\geq 1$  Gyr), and although we miss some of the reddest objects such as 3CR65, we expect this sample adequately to describe the properties of most of the  $z > 0.7$  powerful radio galaxies. Figure 1 shows the spectra of 3CR54, 265 and 343.1, compared to the galaxy companion of 3C245 and the residual noise extracted from the two-dimensional data. The  $S/N$  on the continuum at 4,000 Å is 9, 8, 12, 12 for these four objects respectively. As can be seen in the examples in Fig. 1, despite



FIG. 1 The spectra of (a), the companion galaxy to the quasar 3CR245 at  $z=1.013$  (ref. 10) ( $2 \times 45$  min, ESO/NTT); (b), 3CR343.1,  $z=0.75$  (2 h, CFHT); (c), 3CR265,  $z=0.811$  (1 h, CFHT); (d), 3CR54,  $z=0.827$  (3 h, CFHT); (e), the computed noise of the data reduction. The signal to noise ratios are 12, 12, 8, 9 for the four galaxies respectively. CFHT, Canada-France-Hawaii Telescope; ESO, European Southern Observatory; NTT, New Technology Telescope.

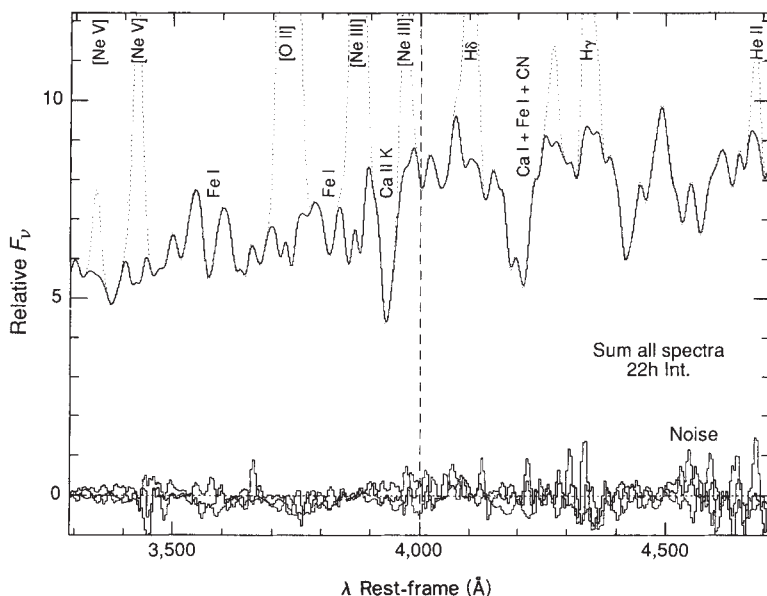


longer integration times none of the spectra of distant radio galaxies show evidence for a 4,000 Å break, although it is clear in the 3CR245 companion spectrum. This absence of signature is also revealed in the sum of all our spectra, for an equivalent exposure time of 22 h (Fig. 2). On this composite spectrum, the  $S/N$  at 4,000 Å is more than 25, and again we do not detect any signature from a break. If the continuum emission from distant radio galaxies had significant contributions from either 'old stars'<sup>1,2</sup>, or from a stellar distribution following the 'range of ages' model<sup>3</sup>, we should expect to detect such a signature (Fig. 2). Moreover, the observed spectral energy distribution is somewhat featureless, although we detect absorption lines in some spectra (Ca and Fe; Fig. 2). A power-law with spectral indices  $\alpha$  ranging from 1.5 to 3.5 (assuming that the flux per frequency unit  $F_\nu = \nu^{-\alpha}$ ) agree well with the observed continuum at ultraviolet wavelengths, which flattens towards redder wavelengths as indicated by near-infrared data (ref. 2). There is no evidence against a single origin for the continuum from

the ultraviolet to the infrared.

We now examine these results in the light of the stellar models already discussed. Any 4,000 Å break amplitude larger than 1.11 would readily be observed in our data with  $S/N \geq 3$ , but is not seen in Fig. 2. Assuming an average 4,000 Å break amplitude between 1.5 and 2.2 (refs 11, 12), our data imply that the contribution from an old stellar population should not exceed 10 to 20%. This contradicts with predictions of both the 'old stars' and the 'range of ages' models. The former model predicts a difference of 1.7 to 2.5 magnitudes between the 'red' and 'blue' populations at 4,000 Å, which is seen in none of the galaxies examined by Rigler *et al.*<sup>2</sup> (see their Fig. 5). The spectra from the 'range of ages' model exhibit 4,000-Å breaks significantly higher than 1.11 (ref. 3), and therefore do not fit our data. Moreover, the shape of our optical spectra does not correspond to the prediction of either model. We therefore believe that models that favour a dominant contribution to the continuum coming from stars older than  $10^8$  years at 4,000 Å do not

FIG. 2 Sum of 10 spectra of distant radio galaxies. They were observed at CFHT with the MARLIN and the MOS-SIS imaging spectrographs and at ESO with the NTT and the imaging spectrograph EMMI, both used in their long slit mode with 2 arcsec slits: 3C54 ( $z=0.827$ ), 3C124 ( $z=1.083$ ), 3C175.1 ( $z=0.919$ ), 3C226 ( $z=0.818$ ), 3C265 ( $z=0.811$ ), 3C280 ( $z=0.996$ ), 3C343.1 ( $z=0.75$ ), 3C352 ( $z=0.806$ ), 3C356 ( $z=1.079$ ), 3C368 ( $z=1.132$ ). The  $S/N$  at 4,000 Å is 25. The 4,000 Å continuum break is not detected, implying a contribution less than 11% from old stars ( $>10^8$  years).



adequately describe the observational data for distant radio galaxies. The model presented by Bithell and Rees<sup>9</sup> predicts spectral energy distributions somewhat closer to what we observe, although again we do not see any of the spectral features expected from a young burst or a short rate of exponentially decaying star formation. Also, Bithell and Rees<sup>9</sup> conclude that these latest models could not account for the small scatter of the  $K-z$  diagram. One can imagine a convolution of several starburst populations hiding a 4,000-Å break (such as in many starburst galaxies), while fitting a steeper slope at ultraviolet wavelengths than at redder wavelengths. Such a model could then be marginally consistent with our data.

The presence of absorption lines in the ultraviolet spectra of distant radio galaxies has been claimed as evidence for the stellar origin of their continua<sup>13</sup>. Such absorption lines, as well as the ones we detect at redder wavelengths, are faint because they are detected in the combination of several spectra, derived from 10 to 22 h of integration. Although these detections definitely prove the presence of stars in distant radio galaxies, they leave open the question of the nature of the continuum. There are indeed several indications of intrinsic correlations between radio and optical properties. Emission line intensities as well as  $K$ -band absolute magnitudes correlate with radio power<sup>5,14</sup>, and a positive correlation between the ultraviolet activity and the radio spectral index has been found<sup>1</sup>. Some of these properties, as well as the radio-optical alignment, can be interpreted as being due to the triggering of star formation by the radio outburst<sup>15</sup>. But this fails to explain several properties of distant radio galaxies, such as the observed line spectra or the fact that some sources have optical emission along and well beyond the radio emission (3CR343.1 is a good example). These properties can also be taken as evidence that the emission line gas is photoionized by an active nucleus<sup>16</sup>. In 3CR368, the extended emission line region shares the properties of the narrow line region of an active galactic nucleus, with the notable exception of the considerably larger spatial extent<sup>7</sup>. Added to the polarization found in several objects of that kind<sup>17,18</sup>, one can argue that a significant fraction of the ultraviolet continuum is light scattered by dust. To date, there is no evidence that this polarization is decreasing towards redder wavelengths, because no dilution with wavelength was found in 3CR265 (ref. 19) and

the observed dilution found in 3CR368 (ref. 17) is probably due to the projection of a galactic foreground M star<sup>7</sup>. Dust-scattering models may reproduce the shape of the spectral energy distribution of distant radio galaxies, and especially the fact that the ultraviolet continuum is steeper than the continuum at longer wavelengths. Polarization measurements at redder wavelengths are obviously crucial.

Our data indicate that the spectral energy distribution of the emission from distant radio galaxies is consistent with much smaller ages than previously believed. Contrary to the current belief, old stars can only contribute a marginal fraction of the observed light from distant radio galaxies. On the other hand, starburst scenarios would be compatible with the age of the radio source itself,  $10^7$  years or less<sup>9</sup>. Therefore, the remaining possible interpretations all seem to indicate that distant radio galaxies might be seen forming their first stars. The small age range would imply that distant radio galaxies, in spite of their scarcity, were the progenitors of an abundant population of galaxies today. Finally, our results lend support to the unification of the different classes<sup>20</sup> of active galactic nuclei. □

Received 11 August 1992; accepted 9 February 1993.

1. Lilly, S. J. *Astrophys. J.* **340**, 77–89 (1989).
2. Rigler, M. A., Lilly, S., Stockton, A., Hammer, F. & Le Fèvre, O. *Astrophys. J.* **385**, 61–82 (1992).
3. Chambers, K. C. & Charlot, S. *Astrophys. J.* **348**, L1–L4 (1990).
4. Spinrad, H. *Publ. astr. Soc. Pacific* **98**, 269–284 (1986).
5. Rawlings, S., Eales, S. A. & Lacy, M. *Mon. Not. R. astr. Soc.* **251**, 17P–21P (1991).
6. Hammer, F. & Le Fèvre, O. *Astrophys. J.* **357**, 38–49 (1990).
7. Hammer, F., Le Fèvre, O. & Proust, D. *Astrophys. J.* **374**, 91–104 (1991).
8. Spinrad, H., Djorgovski, S., Marr, J. & Aguilar, L. *Publ. astr. Soc. Pacific* **97**, 932–961 (1985).
9. Bithell, M. & Rees, M. J. *Mon. Not. R. astr. Soc.* **242**, 570–575 (1990).
10. Le Fèvre, O. & Hammer, F. *astr. Astrophys.* **254**, L29–L32 (1992).
11. Bruzual, A. G. *Astrophys. J.* **273**, 105–127 (1983).
12. Bica, E. *Astr. Astrophys.* **195**, 76–92 (1988).
13. Chambers, K. C. & McCarthy, P. J. *Astrophys. J.* **354**, L9–L12 (1990).
14. Yates, M. G., Miller, L. & Peacock, J. A. *Mon. Not. R. astr. Soc.* **221**, 311–317 (1986).
15. Rees, M. J. *Mon. Not. R. astr. Soc.* **428**, 1P–4P (1989).
16. van Breugel, W. & McCarthy, P. J. in *Proc. ESO Workshop on 'Extranuclear Activity in Galaxies'* (eds Fosbury, R. & Meurs, E.) (ESO, Garching, 1989).
17. di Seregho Alighieri, S., Fosbury, R., Quinn, P. & Tadhunter, C. *Nature* **341**, 307–309 (1989).
18. Tadhunter, C. N., Scarrott, S. M., Draper, P. & Rolph, C. *Mon. Not. R. astr. Soc.* **256**, 53P–58P (1992).
19. Januzzi, B. T. & Elston, R. *Astrophys. J.* **351**, L9–L12 (1991).
20. Heckman, T. M., Chambers, K. C. & Postman, M. *Astrophys. J.* **391**, 39–47 (1992).

ACKNOWLEDGMENTS. We thank C. Boisson and S. Lilly for discussions and the directors of the Canada-France-Hawaii Telescope Corporation for the allocation of discretionary observing times.

## Support for a unified model of radio galaxies and quasars from isotropic [O II] emission

R. Hes\*, P. D. Barthel\* & R. A. E. Fosbury†

\* Kapteyn Astronomical Institute, P.O. Box 800, 9700 AV Groningen, The Netherlands

† Space Telescope-European Coordinating Facility, European Southern Observatory, Karl-Schwarzschild-Strasse 2, 8046 Garching bei München, Germany

UNIFIED models<sup>1–3</sup> of radio-loud quasars and powerful radio galaxies suggest that they are intrinsically similar objects observed from different angles. This can be tested by comparing the isotropically emitted radiation from the spatially extended nebulae surrounding the nuclei; the unified models predict that the intensities of these emissions should be comparable for the two classes of object. But when this prediction was examined<sup>4</sup> for the [O III] 5,007-Å emission line, it was found that quasar [O III] luminosities significantly exceed those of otherwise similar radio galaxies. We have measured the spatially integrated [O II] 3,727-Å emission-line luminosities for a number of quasars and radio galaxies taken from the 3C catalogue. Supplementing our data with values from the literature, we find no systematic difference in the [O II]

luminosities. We argue that this emission is indeed isotropic, and that our results are consistent with the unification hypothesis; the [O III] line, on the other hand, may still have a significant component from the nuclear region, and thus be subject to pronounced anisotropic obscuration.

The striking similarity in the radio properties and cosmological evolution of powerful radio galaxies (PRG) and radio-loud quasars (QSR) has stimulated a debate on the relationship between these classes. The objects might be related in an evolutionary sense, with the bright quasar being a phase in the life of every luminous radio galaxy, or otherwise be more directly connected. In the latter case the question arises of why the QSR are much brighter at optical wavelengths and exhibit broad permitted emission lines with full-width at half-maximum (FWHM) up to  $10^4 \text{ km s}^{-1}$  in their spectra, whereas the permitted lines are narrow ( $\text{FWHM} \leq 10^3 \text{ km s}^{-1}$ ) in PRG.

Unified models have been developed in attempts to explain the differences between several classes of low luminosity active galactic nuclei (such as Seyfert 1 and Seyfert 2 galaxies<sup>1</sup>), through the effects of axis orientation with anisotropic obscuration. Based on these, a radio-loud unified model has been proposed in which narrow-line PRG and QSR are the same objects seen at decreasing angles between the radio axis and the line of sight<sup>2,3</sup>. To explain the difference between the optical continua and the emission line widths, an obscuring torus of dust is postulated: this surrounds the inner region where the broad lines are formed. The PRG are oriented so that this broad



TABLE 1 Samples discussed in the text

| PRG      | z     | Log $P_{178}$ | Log $L[\text{O II}]$ | Ref. | Pair | QSR     | z     | Log $P_{178}$ | Log $L[\text{O II}]$ | Ref. | Pair |
|----------|-------|---------------|----------------------|------|------|---------|-------|---------------|----------------------|------|------|
| 3C300    | 0.270 | 27.78         | 35.67                | 21   | 1    | 3C323.1 | 0.264 | 27.49         | 35.21                | †    | 1    |
| 3C458    | 0.290 | 27.75         | 35.35                | 21   |      | 3C93    | 0.358 | 27.94         | 34.85                | †    | 2    |
| 3C103    | 0.331 | 28.13         | 34.13                | 21   |      | 3C215   | 0.411 | 27.99         | 35.37                | †    | 3    |
| 3C299    | 0.367 | 27.86         | 35.83                | 21   | 2    | 3C47    | 0.425 | 28.38         | 35.55                | 22   | 4    |
| 3C16     | 0.405 | 27.96         | 34.96                | †    |      | 3C455   | 0.543 | 28.25         | 35.73                | †    | 5    |
| 3C142.1  | 0.406 | 28.19         | 35.54                | †    | 3    | 3C147*  | 0.545 | 28.82         | 36.58                | 22   | 6    |
| 3C99     | 0.426 | 28.01         | 36.08                | †    |      | 3C334   | 0.555 | 28.23         | 35.66                | †    |      |
| 3C306.1  | 0.441 | 28.05         | 35.54                | 21   |      | 3C275.1 | 0.557 | 28.44         | 35.79                | †    | 7    |
| 3C341    | 0.448 | 28.02         | 34.92                | 21   |      | 3C345   | 0.594 | 28.18         | 35.45                | 22   | 8    |
| 3C313    | 0.461 | 28.32         | 35.00                | †    | 4    | 3C263   | 0.646 | 28.50         | 36.02                | 22   | 9    |
| 3C435A   | 0.461 | 27.74         | 35.66                | †    |      | 3C37    | 0.670 | 28.27         | 35.39                | 22   | 10   |
| 3C327.1  | 0.463 | 28.38         | 35.21                | †    |      | 3C216*  | 0.670 | 28.61         | 35.54                | 22   |      |
| 3C172    | 0.519 | 28.30         | 35.89                | 21   | 5    | 3C207   | 0.684 | 28.51         | 35.25                | 22   | 11   |
| 3C330    | 0.550 | 28.59         | 36.31                | 21   | 6    | 3C380*  | 0.691 | 29.13         | 35.94                | 22   |      |
| 3C228    | 0.552 | 28.51         | 35.28                | 21   | 7    | 3C254   | 0.734 | 28.75         | 36.22                | 22   | 12   |
| 3C169.1  | 0.633 | 28.18         | 35.93                | 21   | 8    | 3C138*  | 0.759 | 28.69         | 35.66                | 22   | 13   |
| 3C337    | 0.635 | 28.33         | 34.74                | 21   | 9    | 3C39    | 0.765 | 28.35         | 36.18                | 22   | 14   |
| 3C44     | 0.660 | 28.24         | 35.61                | 21   | 10   | 3C175   | 0.768 | 28.76         | 36.18                | 22   | 15   |
| 3C34     | 0.689 | 28.50         | 36.70                | 21   | 11   |         |       |               |                      |      |      |
| 3C441    | 0.707 | 28.50         | 35.42                | 21   | 12   |         |       |               |                      |      |      |
| 3C247    | 0.749 | 28.41         | 36.10                | 21   |      |         |       |               |                      |      |      |
| 3C343.1* | 0.750 | 28.39         | 35.53                | 21   | 13   |         |       |               |                      |      |      |
| 3C277.2  | 0.766 | 28.58         | 36.30                | 21   | 14   |         |       |               |                      |      |      |
| 3C340    | 0.775 | 28.45         | 35.75                | 21   |      |         |       |               |                      |      |      |
| 3C352    | 0.806 | 28.59         | 36.13                | 21   | 15   |         |       |               |                      |      |      |

Data used in this letter. We have adopted  $H_0 = 50 \text{ km s}^{-1} \text{ Mpc}^{-1}$  and  $q_0 = 0.5$ . The radio power  $P_{178}$  is calculated in the rest frame of the objects, using the local spectral index at 178 MHz for the K-correction. The radio power  $\log P_{178}$  is expressed in  $\text{W Hz}^{-1}$ , the line-luminosity  $\log L[\text{O II}]$  in W.

\* Compact steep-spectrum objects.

† R.H., R.A.E.F. and P.D.B., manuscript in preparation.

line region is completely hidden from our view and consequently PRG spectra show no broad lines.

Considerable support exists for the unified model of QSR and PRG. The model seems to be consistent with the linear-size distributions<sup>3</sup>, source counts and luminosity functions<sup>5</sup>. Moreover it indicates solutions to a number of problems, such as the frequent occurrence of superluminal motion in QSR with large radio structures<sup>6</sup> and the apparent deficit of ionizing photons which are required to illuminate the large emission line nebulae surrounding some PRG<sup>7</sup>. More-direct evidence comes from the detection of scattered broad lines in the polarized light of several radio galaxies<sup>8-10</sup>. Further support is provided by the

discovery in the near-infrared of a nuclear point source in the nearest PRG Cygnus A<sup>11</sup>, although the non-detection of scattered broad lines in this object<sup>12</sup> remains to be explained.

This unified model is best tested by comparison of the properties of both classes in isotropically emitted (aspect-independent) radiation. As mentioned, the strong similarity between the extended radio structure of QSR and PRG was a motivation to postulate the model. The outcome of two other comparisons is not in accordance with the simplest version of the model. First, if the PRG harbour a quasar, the large amount of energy radiated in the ultraviolet should be re-radiated by the torus at longer wavelengths (notably the far-infrared) where the dust becomes

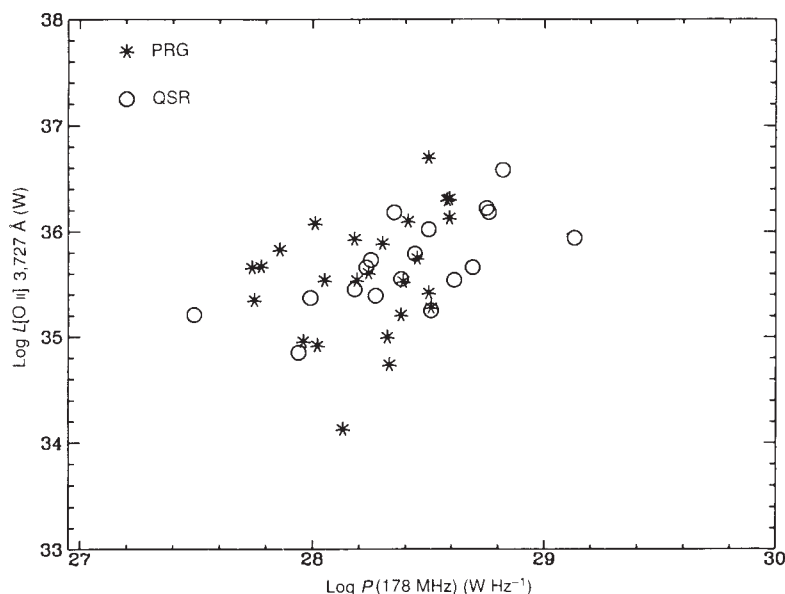


FIG. 1 Distribution of radio galaxies (stars) and quasars (circles). The luminosity in the  $[\text{O II}]$  3,727-Å emission line (in W) is plotted against radio power at 178 MHz in  $\text{W Hz}^{-1}$ .

optically thin. Similar far-infrared luminosities for otherwise comparable QSR and PRG are expected, but careful study of the Infrared Astronomical Satellite database indicates that the QSR in the 3C-catalogue are in the median somewhat brighter at mid and far-infrared wavelengths<sup>13</sup>. This result is not necessarily inconsistent with the unified model, as such a dependence of the infrared flux on orientation is expected if the torus is still optically thick at that frequency<sup>14</sup>. There is also the possibility that a component of the far-infrared radiation is an aspect-dependent (beamed) synchrotron source.

A second argument for intrinsic differences between QSR and PRG has been presented<sup>4</sup>. The unified model arguably predicts that the luminosities of both types in narrow emission lines, which originate from regions beyond the central dust torus, should be similar. It was found however that [O III] 5,007-Å luminosities of QSR were on average 5–10 times higher than those of otherwise similar PRG.

This line may not be the best for this purpose, because a major contribution to the integrated luminosity can be made by nuclear regions on a scale not much larger than the broad line region and so still be partially obscured. We have therefore measured the luminosity of samples of objects in the [O II] 3,727-Å line. This line is more efficiently radiated at low densities because of its lower ionization potential and critical density<sup>15</sup>. The line is therefore believed to originate in a larger volume around the nucleus and to be a better indicator of the isotropic line emission. The unified model predicts that the luminosities of QSR and PRG should be more similar in the [O II] 3,727-Å line than in the [O III] 5,007-Å line. A hint of this tendency is present in the figures shown in refs 16 and 17. In fact, the first high-redshift radio galaxies were discovered thanks to their high, quasar-like [O II] 3,727-Å luminosities<sup>18</sup>.

Because the nebulosities emitting the [O II] 3,727-Å line can be extended in both PRG and QSR, we carried out a project to obtain narrow-band images of samples of objects in the redshifted [O II] 3,727-Å line. An advantage over spectroscopy is that no emission is missed because of slit orientation and limited slit width. The observations were performed with the European Southern Observatory (ESO) 3.6-m telescope, equipped with the ESO Faint Object Spectrograph and Camera. (R.H., R.A.E.F. and P.D.B., manuscript in preparation.)

Objects were selected from an equatorial subsample of the 3C compilation<sup>19</sup>. At the time of our measurements this was the only nearly completely identified catalogue of objects selected at low radiofrequency (178 MHz). There is ample evidence that steep-spectrum low-frequency emission is unaffected by beaming<sup>20</sup>, so that the selection is based on an isotropic property and hence no bias is introduced towards sources close to the line of sight. Our sample objects satisfy the criterion that the radio emission at 178 MHz is dominated by steep-spectrum emission, the quasar 3C345 being the only marginal case. We chose samples of PRG and QSR with redshift in the range  $0.25 < z < 0.8$ , as in ref. 4.

Table 1 lists the total [O II] 3,727-Å luminosities of our sample sources. We have added luminosities of all PRG within the same redshift bin with reliably calibrated [O II] 3,727-Å imaging from ref. 21 and also a number of QSR [O II] luminosities from recent spectroscopy<sup>22</sup>. Figure 1 shows the distribution of sources in the plane defined by the extended radio power  $P_{178}$  and the [O II] 3,727-Å line luminosity.

Figure 1 shows no separation between the classes in their [O II] 3,727-Å luminosities  $L[\text{O II}]$ : at a given radio luminosity there is no tendency for the QSR to have a higher  $L[\text{O II}]$  than the PRG. We find a median value of  $\log L[\text{O II}]$  of 35.61 W for the PRG and a median of 35.66 W for the QSR. The trend visible in the distributions is the increase of  $L[\text{O II}]$  with radio power, which appears to be the same for different types of active galactic nuclei spanning a large range in radio power (refs 17, 23, 24).

From the samples we have constructed 15 pairs of PRG and QSR that are matched in redshift within 0.05 and in radio luminosity within a factor of two (see Table 1). In cases where there was more than one PRG to match a QSR we took the galaxy nearest to the quasar in the  $P_{178}$ - $z$  plane. By matching the samples this closely, the possible dependence of  $L[\text{O II}]$  on evolution and the known dependence on the power of the extended radio emission<sup>4,17</sup> are eliminated. We now find for the PRG a mean  $\log L[\text{O II}]$  of 35.73 W and for the QSR of 35.70 W. We have performed a Wilcoxon matched pairs signed rank test and find that we cannot reject the hypothesis that the samples are drawn from an identical parent population.

This result is clearly consistent with the prediction of the unified model. We assume that the extended narrow line emission is photoionized by the nuclear source, which is justifiable for both classes<sup>25,26</sup>. In this case the size of the dust torus is constrained by the volumes from which the [O II] 3,727-Å and [O III] 5,007-Å lines are emitted. Assuming a range of filling-factors<sup>21,25</sup> from  $10^{-3}$  to  $10^{-6}$ , a temperature  $T = 10^4$  K and the typical densities at which [O II] 3,727-Å is emitted, the observed mean luminosity  $\log L[\text{O II}] = 35.7$  W implies a minimum emitting volume with radius  $r = 0.5$ –5 kpc. A recent estimate of the outer torus size is  $r < 800$  pc, based on high-resolution imaging of the PRG Cygnus A<sup>27</sup>. Although coarse, these estimates seem to justify the notion that [O II] 3,727-Å is a better tracer of the isotropic emission. This is also borne out by the fact that the measured sizes of the [O II] nebulae are up to 100 kpc. Finally we note that if the [O III] 5,007-Å is indeed partially obscured by the dust torus, part of the emission may show up weakly in the reflected polarized spectrum.

We conclude that the previously reported difference<sup>4</sup> in the emission-line properties of QSR and PRG is ruled out as a powerful argument against the unified picture, and that new tests have to be devised to refine or reject the model. □

Received 7 December 1992; accepted 9 February 1993.

- Antonucci, R. R. J. & Miller, J. S. *Astrophys. J.* **297**, 621–632 (1985).
- Scheuer, P. A. G. in *Superluminal Radio Sources* (eds Zensus, J. A. & Pearson, T. J.) 104–113 (Cambridge Univ. Press, 1987).
- Barthel, P. D. *Astrophys. J.* **336**, 606–611 (1989).
- Jackson, N. & Browne, I. W. A. *Nature* **343**, 43–45 (1990).
- Padovani, P. & Urry, C. M. *Astrophys. J.* **387**, 449–457 (1992).
- Barthel, P. D., Hooimeyer, J. R., Schilizzi, R. T., Miley, G. K. & Preuss, E. *Astrophys. J.* **336**, 601–605 (1989).
- Tadhunter, C. N. in *New Windows to the Universe* Vol. 2 (eds Sanchez F. & Vasquez, M.) 175–192 (Cambridge Univ. Press, 1990).
- Antonucci, R. *Nature* **299**, 605–606 (1982).
- Goodrich, R. W. & Cohen, M. H. *Astrophys. J.* **391**, 623–628 (1992).
- Antonucci, R. & Barvainis, R. *Astrophys. J.* **363**, L17–L20 (1990).
- Djorgovski S., Weir, N., Matthews, K. & Graham, J. R. *Astrophys. J.* **372**, L67–L70 (1991).
- Jackson, N. in *Physics of Active Galactic Nuclei* (eds Duschl, W. J. & Wagner, S. J.) 612–617 (Springer, Berlin, 1992).
- Heckman, T. M., Chambers, K. C. & Postman, M. *Astrophys. J.* **391**, 39–47 (1992).
- Pier, E. & Krolik, J. *Astrophys. J.* **401**, 99–109 (1992).
- Osterbrock, D. *Astrophysics of Gaseous Nebulae and Active Galactic Nuclei* (University Science Books, Mill Valley, 1989).
- Browne, I. W. A. & Jackson, N. in *Physics of Active Galactic Nuclei* (eds Duschl W. J. & Wagner, S. J.) 618–635 (Springer, Berlin, 1992).
- Rawlings, S. & Saunders, R. *Nature* **349**, 138–140 (1991).
- Spinrad, H. *Publs. astr. Soc. Pacif.* **98**, 269–284 (1986).
- Spinrad, H., Djorgovski, S., Marr, J. & Aguilar, L. *Publs. Astr. Soc. Pacif.* **97**, 932–961 (1985).
- Hough, D. H. & Readhead, A. C. S. *Astr. J.* **98**, 1208–1225 (1989).
- McCarthy, P. J. thesis, Univ. California, Berkeley (1988).
- Jackson, N. & Browne, I. W. A. *Mon. Not. R. astr. Soc.* **250**, 414–421 (1991).
- Saunders, R., Baldwin, J. E., Rawlings, S., Warner, P. J. & Miller, L. *Mon. Not. R. astr. Soc.* **238**, 777–790 (1989).
- Baum, S. A. & Heckman, T. M. *Astrophys. J.* **336**, 702–771 (1989).
- Stockton, A. & MacKenty, J. W. *Astrophys. J.* **316**, 584–596 (1987).
- Robinson, A., Binette, L., Fosbury, R. A. E. & Tadhunter, C. N. *Mon. Not. R. astr. Soc.* **227**, 97–114 (1987).
- Vestergaard, M. & Barthel, P. D. *Astr. J.* **105**, 456 (1993).

ACKNOWLEDGEMENTS. We thank R. Antonucci for comments. R.H. thanks the European Southern Observatory for support and P.D.B. acknowledges support from the Royal Netherlands Academy of Arts and Sciences. R.A.E.F. is affiliated to the Astrophysics Division, Space Science Department, European Space Agency.



# Transition between crack patterns in quenched glass plates

A. Yuse & M. Sano

Research Institute of Electrical Communication, Tohoku University, Sendai 980, Japan

THE study of fracture is an old topic<sup>1</sup>, but only recently has an understanding begun to emerge of crack formation, propagation and morphology (which is often fractal)<sup>2-8</sup>. When a brittle material such as glass is broken under tensile stress<sup>9</sup>, the cracks have a complicated morphology<sup>10</sup>. Fineberg *et al.*<sup>11</sup> showed that this process may be caused by a dynamic instability, whereby the speed of crack propagation increases until it approaches the speed of sound: at this point, complex structures appear. But crack morphology in quasi-static fracture, where the speed of the crack tip is much smaller than the speed of sound, can also exhibit marked changes<sup>12</sup>. Here we present studies of crack propagation in glass plates caused by sudden but carefully controlled cooling. We observe a transition from straight to regular, wavy cracks as the tip speed increases. The scaling behaviour of an appropriately defined relaxation time suggests that this transition is a Hopf bifurcation<sup>13</sup>, like those seen in a variety of other nonlinear systems. At still higher speeds, the oscillatory cracks split into first two and then four or more branches.

The experimental apparatus is shown in Fig. 1. Thermal stress is applied along a glass plate by placing it between two heat reservoirs, a heater (D) and a water bath (E), at different temperatures. The highest thermal stress occurs in the part of the sample lying between the reservoirs. The sample is then moved at a velocity  $v$ . If the thermal stress is large enough to

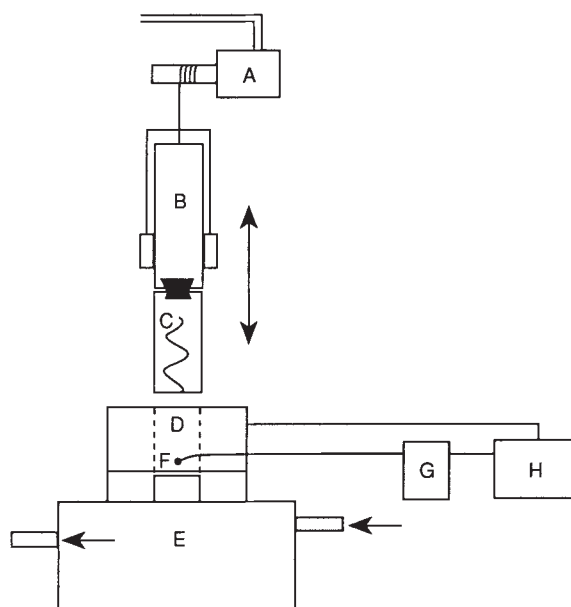


FIG. 1 Experimental arrangement. D is a heater (controlled to an accuracy of  $\pm 0.2$  K) consisting of an aluminium block with a slot of depth 0.6 mm. E is a cooled water bath (controlled to an accuracy of  $\pm 0.1$  K). The temperature difference between them is typically 70 K and the separation is  $10.5 \pm 0.5$  mm. The glass plate (C) is attached to the linear slide (B), which is suspended on a stainless steel wire wound on a brass cylinder driven by the computer-controlled stepping motor (A). F is a temperature sensor, G and H are temperature controllers.

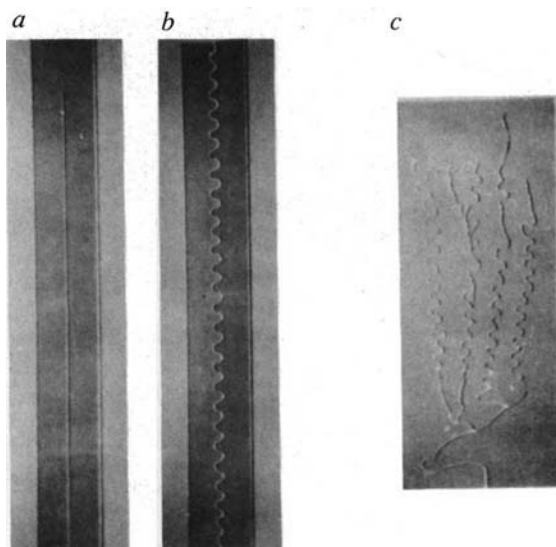


FIG. 2 Typical examples of three different types of cracks: a, a straight crack, b, oscillating crack, c, branched crack. Sample size (a and b)  $10 \times 100 \times 0.11$  mm<sup>3</sup>, (c)  $24 \times 60 \times 0.13$  mm<sup>3</sup>.

create new surface, an existing crack can propagate. Its tip will move at a vertical velocity of  $-v$  to maintain the critical stress. The typical temperature difference  $\Delta T$  between reservoirs D and E is 70 K, and their separation is  $10.5 \pm 0.5$  mm. The velocity  $v$  can be varied between 0.5 and 50 mm s<sup>-1</sup> with an accuracy of  $\pm 1.5\%$ .

A glass sample was first inserted through the slot of the heater, warmed for a few minutes to equilibrate, then lowered into the water bath. To minimize heat capacity and thermal diffusion time, and to make the crack patterns quasi-two-dimensional, we used thin cover glass ( $d = 0.13$  mm thick) of different aspect ratios. The thermal conductivity of the sample is  $\lambda = 2.0 \times 10^{-3}$  cal cm<sup>-1</sup> s<sup>-1</sup> K<sup>-1</sup>. The density is  $\rho = 2.53$  g cm<sup>-3</sup> and the specific heat is  $C_p = 0.17$  cal g<sup>-1</sup> K<sup>-1</sup>, so the thermal diffusivity is  $\kappa = 4.7 \times 10^{-3}$  cm<sup>2</sup> s<sup>-1</sup>. From the typical velocity  $v$  at which oscillation sets in, we obtain the thermal diffusion length  $l_{th} = \kappa/v \approx 0.6$  mm, 10 times larger than  $d/2$ . Thus the temperature profile along the plate in the gap is an exponentially decaying function.

When the sample begins to cool on entering the bath, many cracks can start from the edge<sup>12</sup>. The average number of propagating cracks is nearly proportional to  $\Delta T$  and sample width  $L$ . The number of cracks affects the results. We therefore made a small vertical notch in the centre of the bottom of the glass plate to ensure the nucleation of a single propagating crack. The crack morphology depends on  $\Delta T$  and  $v$ , as shown in Fig. 2. There are three types: straight (Fig. 2a), oscillating (Fig. 2b) and branched (Fig. 2c).

Oscillating cracks (Fig. 2b) show a transient stage, after which the amplitude of oscillation saturates. Figure 2c shows a crack which branched in two, then in four: further branching was also observed. Figure 3 shows a phase diagram for the three types of crack using glass of dimensions  $24 \times 64 \times 0.13$  mm<sup>3</sup>. The phase diagram is similar for thicker samples, except that crack patterns with a network structure occurred for glass at a thickness of 1.00 mm, but not at 0.13 mm. Such network structures are omitted from the phase diagram. For small  $\Delta T$ , cracks do not grow. The solid line in Fig. 3 represents the boundary  $v_c(\Delta T)$  above which a crack oscillates. The dotted line denotes the threshold for the appearance of branched cracks. Crack surfaces were smooth (mirror-like) for all cases. From the typical  $\Delta T$  for crack onset, 100 K, we can estimate the breaking stress as  $\sigma_{th} = \alpha E \Delta T / 2(1 - \epsilon)$ , where  $E$  is the Young's modulus ( $7.23 \times 10^5$  kg cm<sup>-2</sup> for the sample),  $\epsilon$  is the Poisson's ratio (0.23), and

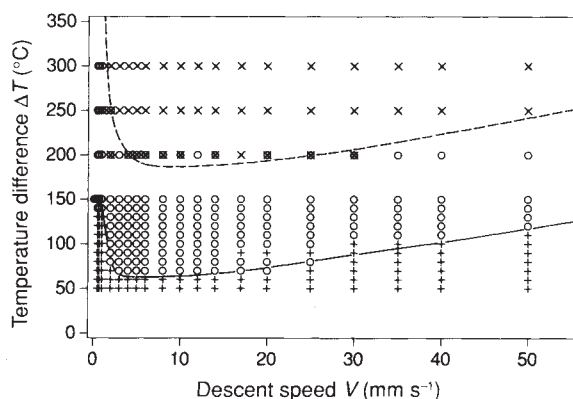


FIG. 3 Morphology phase diagram in  $v$  against  $\Delta T$ . +, straight crack; O, oscillating crack; x, branched crack. The solid line is the boundary above which a crack oscillates, and the dotted line is the threshold for the appearance of branched cracks.

$\alpha$  is the thermal expansion rate ( $0.77 \times 10^{-5} \text{ K}^{-1}$ ). The estimated value  $\sigma_{th} \approx 400 \text{ kg cm}^{-2}$  is close to the nominal breaking stress of the sample<sup>14</sup>.

When the velocity is a little beyond the onset of oscillation, the crack shape is almost sinusoidal (Fig. 4a). As  $v$  or  $\Delta T$  is increased, the shape resembles a series of semicircles (Fig. 4b) and finally becomes asymmetric in the direction of propagation. At very high speeds the crack propagates in an almost horizontal direction until it stops at some distance from the side, then grows upwards forming a cusp (Fig. 4c).

The transition from straight to oscillating cracks (given by the solid line in Fig. 3) was examined in detail. We fixed  $\Delta T$  and changed  $v$  in small steps around  $v_c(\Delta T)$ . For some initial conditions, a crack can exhibit damped oscillation for  $v$  a little below  $v_c(\Delta T)$ . The damped oscillation occurs if the heated edge of the glass plate is suddenly lowered into the water, but not if the edge is inserted into the water before heating. The crack patterns can be fitted well by the function  $A \exp(-x/l) \cos(kx + x_0)$ , where the decay length  $l$  depends on  $v$ . The appropriate variable is the relaxation time,  $\tau = l/v$ . In Fig. 5a we plot the scaling of  $\tau = l/v$  as a function of  $(v - v_c)$ .  $\tau$  diverges when  $v$  approaches  $v_c$  as  $\tau \sim |v - v_c|^{-0.96 \pm 0.09}$  ( $l$  itself diverges as  $l \sim |v - v_c|^{-1.05}$ ), implying that the transition is a supercritical Hopf bifurcation<sup>13</sup>.

To understand the scaling, we assume by symmetry that the amplitude of the oscillation obeys a Ginzburg-Landau equation:

$$\partial_t A = \varepsilon A - A^3 \quad (1)$$

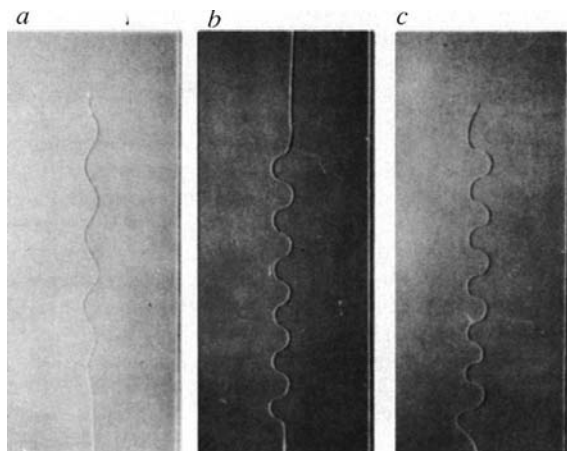


FIG. 4 Development of an oscillating crack. a, Sinusoidal oscillation; b, saturated oscillation; c, nonlinear oscillation. (Sample size  $24 \times 64 \times 0.13 \text{ mm}^3$ .)

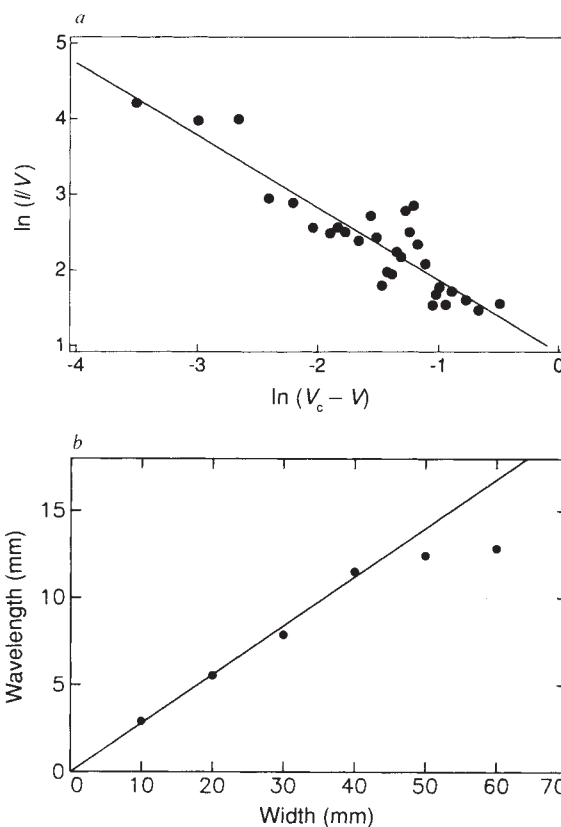


FIG. 5 a, Scaling relation between the decay time  $\tau = l/v$  and the reduced velocity  $|v - v_c|$ . The straight line is the best fit with slope  $-0.96$ . b, Scaling relation between the wavelength of oscillation  $\lambda$  and the width of the sample  $L$ . The best fit for the data with  $L \leq 40 \text{ mm}$  is  $\lambda \sim 0.28 L^{1.0}$ .

where  $A$  denotes the amplitude of oscillation and  $\varepsilon$  is the reduced control parameter,  $\varepsilon = (v - v_c)/v_c$ . We use a co-moving coordinate  $x = x_0 + vt$  for the  $x$  position of the tip. Making use of the relation  $\partial_t A = v \partial_x A$ , we obtain

$$\partial_x A = (\varepsilon/v)A - (1/v)A^3 \quad (2)$$

For small negative  $\varepsilon$  and  $A$ , we can neglect the  $A^3$  term and obtain a damped oscillating solution,  $A = A_0 \exp(-x|\varepsilon|/v)$ , implying that  $\tau \sim \varepsilon^{-1.0}$ .

To see what determines the length scale of an oscillating crack, we measured the scaling of wavelength,  $\lambda$ , for sample width,  $L$ , ranging from 10 mm to 60 mm. As  $\lambda$  depends weakly on  $v$ , we fixed  $v$  and chose  $\Delta T$ , so that  $v \approx v_c(\Delta T, L)$ . When  $L \leq 40 \text{ mm}$ ,  $\lambda$  scales as  $0.28 L^{1.0}$  (see Fig. 5b), showing that the tip of the crack is strongly influenced by the lateral boundary conditions. In an infinite or laterally periodic sample, instead of oscillating, the crack will propagate at a fixed angle. In fact, in cylindrical samples, we observed helical cracks with pitch depending on  $v$  and  $\Delta T$ . The oscillation may arise from the competition between two effects, the tendency to lengthen the crack to release free energy, and the restoring force driving the crack tip toward the centre away from the lateral edges. Any theory must include  $L$  and the lateral boundary condition to describe the instability in quasi-static fracture. □

Received 21 December 1992; accepted 5 February 1993.

1. Griffith, A. A. *Phil. Trans. R. Soc. A* **221**, 163-198 (1920).
2. Herrmann, H. J. & Roux, S. (eds) *Statistical Models for the Fracture of Disordering Media* (North-Holland, Amsterdam, 1990).
3. Charnet, J. C., Roux, S. & Guyon, E. (eds) *Disorder and Fracture* (Plenum, New York, 1990).
4. Mandelbrot, B. B. & Paullay, D. E. *Nature* **308**, 721-722 (1984).
5. Takayasu, H. *Prog. theor. Phys.* **74**, 1343-1345 (1985).
6. Skjeltop, A. T. & Meakin, P. *Nature* **335**, 424-426 (1988).



7. Herrmann, H. J., Kertesz, J. & de Arcangelis, L. *Europhys. Lett.* **10**, 147–152 (1989).
8. Mori, Y., Kaneko, K. & Wadachi, M. *J. phys. Soc. Japan*, **60**, 1591–1599 (1991).
9. Yoffe, E. H. *Phil. Mag.* **42**, 739–750 (1951).
10. Ravi-Chandar, K. & Knauss, W. G. *Int. J. Fracture* **25**, 247–252 (1984); **26**, 65–80 (1984); **26**, 141–154 (1984); **26**, 189–200 (1984).
11. Fineberg, J., Gross, S. P., Marder, M. & Swinney, H. L. *Phys. Rev. Lett.* **67**, 457–460 (1991).
12. Hirata, M. *Scient. Pap. Inst. phys. chem. Res. (Japan)* **16**, 172–195 (1931).
13. Guckenheimer, J. & Holmes, P. *Nonlinear Oscillations and Bifurcation of Vector Fields* (Springer, Berlin, 1983).
14. Stevels, J. M. in *Handbuch der Physik* Vol. 13 (ed. Flugge, S.) (Springer, Berlin, 1962).

ACKNOWLEDGEMENTS. We thank Y. Sawada, F. Graner, J. A. Glazier, and E. Moses for comments and discussions.

## Role of sulphur photochemistry in tropical ozone changes after the eruption of Mount Pinatubo

S. Bekki, R. Toumi & J. A. Pyle

Centre for Atmospheric Science, Department of Chemistry,  
University of Cambridge, Lensfield Road, Cambridge CB2 1EW, UK

RECENT observations suggest that the eruption of Mount Pinatubo in June 1991 has had a considerable effect on ozone concentrations in the tropical stratosphere (refs 1, 2, and J. W. Waters, personal communication). Although stratospheric ozone losses following volcanic eruptions are generally attributed to the presence of sulphate aerosol<sup>3–7</sup>, we present model calculations which demonstrate that gas-phase sulphur chemistry may have played a part in the tropical ozone perturbations that followed the Pinatubo eruption. We find that in the first month or so after the eruption, the large amount of SO<sub>2</sub> injected into the tropical atmosphere catalyses mid-stratospheric ozone production. On the other hand, the SO<sub>2</sub> cloud absorbs solar radiation, thereby reducing the rate of O<sub>2</sub> photolysis (and hence of ozone production) below it. These two effects cancel each other out at an altitude of about 25 kilometres. After one or two months, most of the SO<sub>2</sub> has been oxidized to sulphate; the efficiency of these two mechanisms then becomes negligible (although ozone remains perturbed in the lower stratosphere because of its long photochemical lifetime in this region). The model features show good agreement with initial ozone measurements following the eruption, including both the mid-altitude switch from ozone loss to ozone gain<sup>1</sup>, and the increase and subsequent decrease in the total ozone column<sup>2,7</sup>.

Ozonesonde measurements averaged over August and September 1991 show that the concentration of O<sub>3</sub> had substantially decreased in the tropical lower stratosphere compared with measurements averaged over the quarter preceding the Mount Pinatubo eruption<sup>1</sup>. The maximum O<sub>3</sub> reduction (18–20%) was observed at an altitude around 24–25 km. Interestingly, the data also suggest that O<sub>3</sub> concentrations had increased by 5–10% above 28 km. Analysis of the Total Ozone Mapping Spectrometer (TOMS) data indicates local increases in the tropical O<sub>3</sub> column of 2% in July, but decreases of 1% in August and 3–4% in September (Table 1)<sup>2</sup>. When zonally averaged, TOMS O<sub>3</sub> changes do not exceed 2–4% (ref. 8). Ozone measurements obtained by the Microwave Limb Sounder (MLS) experiment in September (three months after the eruption), when compared to climatological values, suggest O<sub>3</sub> reductions in the tropical lower stratosphere, but no significant O<sub>3</sub> increases around 30 km (J. W. Waters, personal communication). Some of these O<sub>3</sub> datasets<sup>1</sup> have been analysed without completely eliminating the dynamical variations related to the quasi-biennial oscillation<sup>2,8</sup>. Therefore the observed O<sub>3</sub> changes cannot be entirely attributed to the Mount Pinatubo eruption.

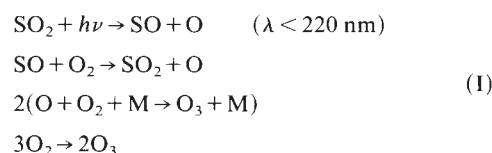
A number of possible causes have been advanced to explain stratospheric O<sub>3</sub> reductions following large volcanic eruptions (although these may not all apply to the tropics). The mechanisms currently proposed attribute O<sub>3</sub> losses to perturbations

induced by large increases in the stratospheric sulphate aerosol loading: these include enhanced heterogeneous chemistry on aerosols<sup>3,4</sup>, enhanced scattering from the aerosols<sup>5</sup> and radiative heating by the aerosols causing enhanced upwelling<sup>6,7</sup>.

Before the volcanic cloud was dispersed and SO<sub>2</sub> converted to sulphate aerosols, O<sub>3</sub> perturbations caused by large concentrations of SO<sub>2</sub> in the volcanic cloud may have occurred. The eruption of Mount Pinatubo injected between 15 and 30 Mt of SO<sub>2</sub> into the stratosphere<sup>6,9,10</sup>; SO<sub>2</sub> was oxidized and converted to sulphate aerosols in one or two months.

In the stratosphere, the rate-limiting step in the oxidation chain leading to the conversion of SO<sub>2</sub> to H<sub>2</sub>SO<sub>4</sub> is believed to be the reaction of SO<sub>2</sub> with OH. The HO<sub>x</sub> (=OH + HO<sub>2</sub>) lost is subsequently regenerated, so that the net effect of the oxidation sequence on HO<sub>x</sub> (and hence on O<sub>3</sub>) is null<sup>11,12</sup>. The end oxidation product, H<sub>2</sub>SO<sub>4</sub>, forms new sulphate particles or condenses on pre-existing particles.

Ozone production catalysed by SO<sub>2</sub> is also possible through the following sequence<sup>13</sup>



Sulphur dioxide may also affect O<sub>3</sub> by absorption of radiation. SO<sub>2</sub> absorbs strongly between 180 and 235 nm, weakly between 260 to 340 nm and very weakly between 340 to 390 nm<sup>14</sup>. Therefore SO<sub>2</sub> has the potential to reduce the transmission of the solar flux, reducing the photolysis rates of key species such as O<sub>2</sub> and hence reducing the rate of ozone production.

In order to assess the importance of the two mechanisms mentioned above, we use a zonally averaged chemical-radiative-dynamical model<sup>15,16</sup> of the middle atmosphere. The chemical scheme combines what are thought to be the most important O<sub>x</sub>, NO<sub>x</sub>, HO<sub>x</sub>, ClO<sub>x</sub>, BrO<sub>x</sub>, CHO<sub>x</sub> and SO<sub>x</sub> (ref. 17) reactions with the recommended kinetic data<sup>18</sup>. The values of SO<sub>2</sub> absorption cross-sections adopted in this work are taken from ref. 19 for the 180–235 nm region and from ref. 14 for the 260–390 nm region.

Model integrations cover the period from June to September 1991. The volcanic eruption of Mount Pinatubo is simulated by injecting 20 Mt of SO<sub>2</sub> into one model column (9.5° wide, centred at 9.47° N) over the altitude range of 21 to 28 km from the 12 to 15 June<sup>9,10,20</sup>. Certain simplifications are made; an instantaneous mixing of the SO<sub>2</sub> emissions into the model grid boxes is assumed, subgrid processes such as plume chemistry are

TABLE 1 Changes in the tropical ozone column (Dobson units)

|             | July | August | September |
|-------------|------|--------|-----------|
| Simulation* | +5.4 | −0.3   | −2.5      |
| TOMS†       | +5   | −2     | −8        |
| Ozonesonde‡ |      | −14    | −18       |

\* O<sub>3</sub> column changes are derived by comparison to the control run (no SO<sub>2</sub> injection) and averaged over the 14° S–14° N latitude band.

† O<sub>3</sub> column changes correspond to changes in the minimum O<sub>3</sub> column from the TOMS daily map in the 12° S–12° N latitude band. The changes presented here were derived by comparison to the O<sub>3</sub> column map in June, because the climatology indicates that the tropical ozone column stays relatively constant from June to September<sup>2</sup>. It is worth noting that tropical ozone column normally drops by 15–20 Dobson units from September to December<sup>2</sup> and that therefore this procedure is no longer applicable after September.

‡ O<sub>3</sub> column changes are derived by comparison to the O<sub>3</sub> column averaged over the quarter preceding the Mt Pinatubo eruption (this is April, May and June)<sup>1,7</sup>. No measurements by ozonesondes could be performed above 31 km, so that only a part of the region showing O<sub>3</sub> enhancement was scanned. Therefore these O<sub>3</sub> column changes are likely to be upper limits.

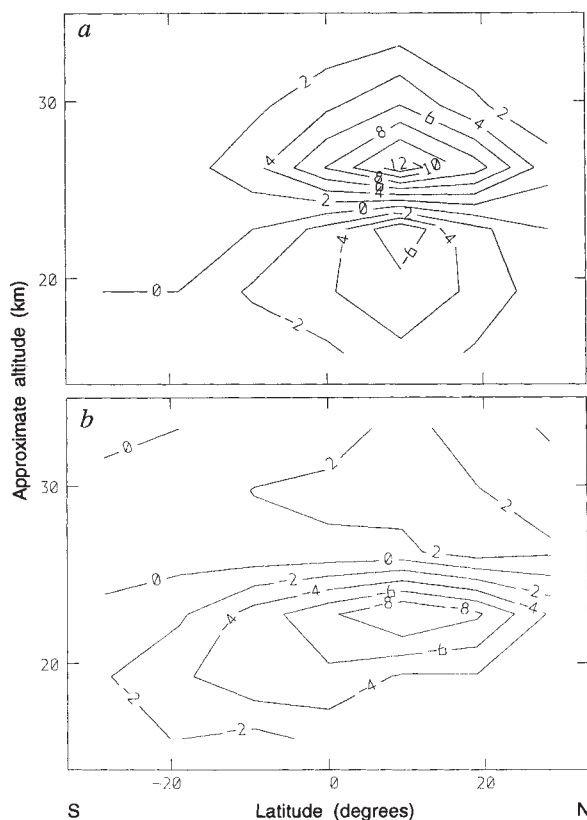


FIG. 1 Modelled percentage changes in the  $O_3$  concentration as a function of latitude and altitude in July (a) and September (b); comparison is between the simulation including  $SO_2$  injection and a control run with no  $SO_2$  injection.

neglected and no heterogeneous chemistry or aerosol radiative heating is included.

We focus our attention on the three months following the eruption, when the  $SO_2$ -induced perturbations are expected to be the most significant. Figures 1a and b shows the calculated changes in  $O_3$  concentration in July and September (one and three months after the eruption, respectively) due to the  $SO_2$  injections. The changes in  $O_3$  concentrations are predominantly confined to the latitude band of injections in July. Ozone concentrations increase in the mid-stratosphere (the top region of the volcanic cloud), by up to 12% at around 27 km. In the tropical lower stratosphere, however,  $O_3$  is reduced by up to 7%. By September, the initial  $SO_2$  cloud has spread throughout the tropics. Ozone increases ( $\sim 2\%$ ) in the mid-stratosphere are now much smaller than in July, but  $O_3$  in the tropical lower stratosphere remains low with reductions of up to 8%.

Ozone increases calculated in the mid-stratosphere are caused by the large concentrations of  $SO_2$  ( $\sim 200$  p.p.b. at 25 km in July) present in the initial volcanic cloud, which catalyse  $O_3$  production through cycle I. Production of  $O_3$  by this mechanism becomes comparable to the  $O_3$  production by photolysis of  $O_2$  in the  $SO_2$  cloud in July (Fig. 2a). The calculated increases in  $O_3$  are, however, lower than anticipated from the additional  $SO_2$ -catalysed production of  $O_3$  alone. This can be explained by the decrease in the solar flux (caused by  $SO_2$  absorption) which reduces the rate of  $O_2$  photolysis and hence of  $O_3$  production. In the lower stratosphere in July (Fig. 3a)  $O_2$  photolysis is reduced by up to 50%. This  $O_3$ -reducing mechanism competes with the  $O_3$ -producing cycle I throughout the stratosphere. Because the rate of cycle I depends on the local concentration of  $SO_2$ , it shows a maximum in the  $SO_2$  cloud. Conversely, the reduction in  $O_2$  photolysis is a maximum below the cloud because of the large  $SO_2$  column. The opposite effects on  $O_3$  of

the two processes compensate each other around 25 km in the model.

The rates of these processes depend on the amount of  $SO_2$  present in the stratosphere. The model predicts a lifetime of about 45 days for  $SO_2$ , which is consistent with the observed  $SO_2$  decay time ( $\sim 35$  days<sup>9</sup>). As most of the  $SO_2$  cloud is converted to sulphate by September, the perturbations of the photochemical balance of  $O_3$  are substantially reduced at this time compared with July;  $O_3$  production through cycle I now represents less than 10% of the  $O_3$  production in the mid-stratosphere (Fig. 2b), and oxygen photolysis is reduced only by about 10% in the lower stratosphere (Fig. 3b). The recovery of the normal photochemical balance for  $O_3$  is evident in the mid-stratosphere with  $O_3$  increases much lower in September ( $\sim 2\%$ ) compared to July ( $\sim 12\%$ ). In the lower stratosphere, however,  $O_3$  is still 8% lower than the control run values because of the long lifetime for  $O_3$  in this region. The photochemical replacement time for  $O_3$  is of the order of several months in the tropical lower stratosphere compared to only about a week in the mid-stratosphere. This mechanism explains why measurements<sup>1</sup> made during the quarter following the eruption indicate significant  $O_3$  increases in the mid-stratosphere, whereas observations<sup>3</sup> (J. W. Waters, personal communication) made after this period do not show any significant changes in that region. The model calculations also capture an important feature, the switching from positive to negative anomaly in the  $O_3$  column shown by the TOMS data (Table 1). None of the aerosol-induced mechanisms currently proposed can account for this feature. The modelled  $O_3$  column reductions in August and September are smaller than observed<sup>2</sup>, suggesting that  $O_3$ -destroying mechanisms (probably related to the large increase in sulphate aerosol loading) may be missing in the simulations. No firm conclusions can be drawn, however, because the analysis of the TOMS data<sup>2</sup> focuses on the largest local changes whereas

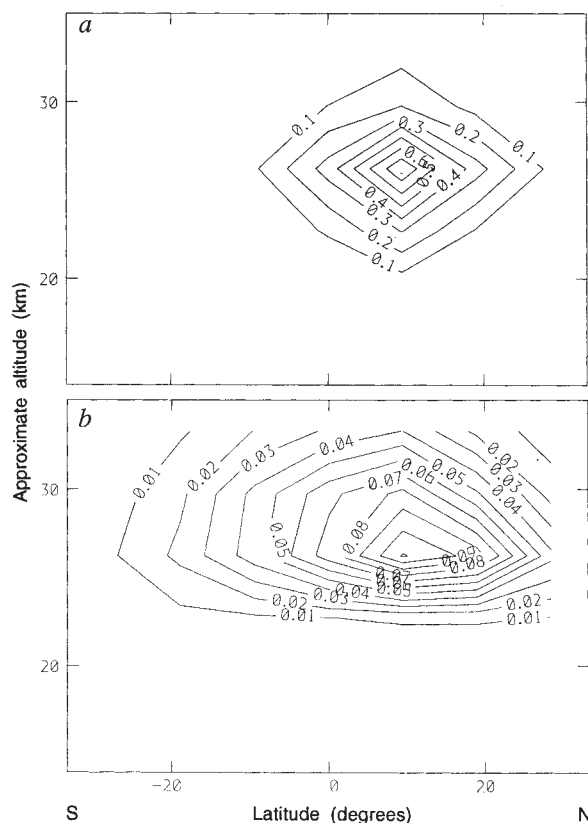


FIG. 2 Modelled ratio of the  $O_3$  production from cycle I to the  $O_3$  production from  $O_2$  photolysis ( $J_{SO_2}/J_{O_2}$ ) as a function of latitude and altitude in July (a) and September (b).



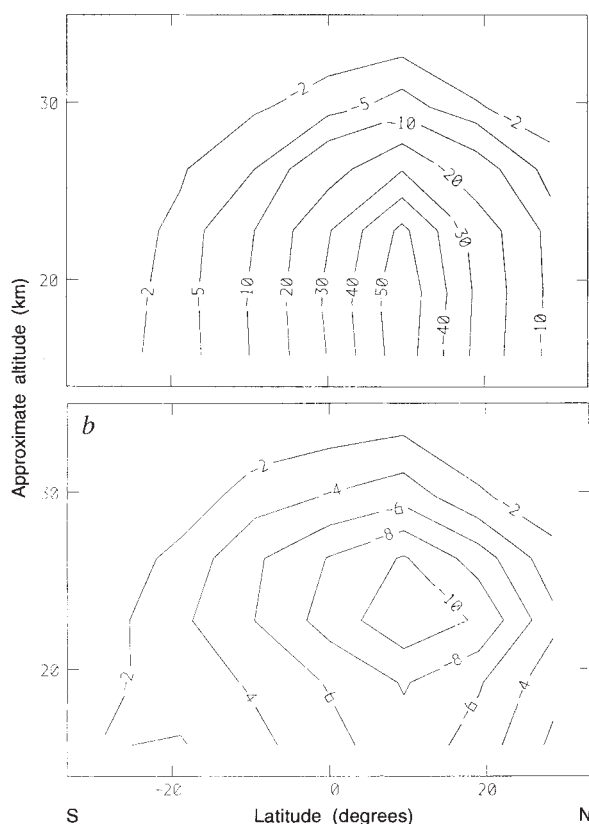


FIG. 3 Modelled percentage changes in the  $O_2$  photolysis as a function of latitude and altitude in July (a) and September (b); comparison is between the simulation including  $SO_2$  injection and a control run with no  $SO_2$  injection.

the model calculations correspond to zonally averaged quantities.

There are other discrepancies between the calculations and observations. Measurements suggest<sup>1</sup> that the crossing over from production to destruction of  $O_3$  occurs at an altitude of 28 km. Model predictions on the other hand indicate an altitude of about 25 km. This may be due to the coarse resolution of the model ( $\sim 3.5$  km vertically) and a number of uncertainties remaining in the model calculations, such as the initial distribution of the volcanic cloud and the amount of  $SO_2$  injected. In addition this version of the model does not include the quasi-biennial oscillation, an important feature of the tropical stratospheric circulation<sup>21</sup>. One consequence of this is that although the observations suggest that the cloud was centred on the Equator, the model simulates a cloud at about  $10^\circ$  N, because it does not allow for rapid transport of the cloud into the tropics and Southern Hemisphere. Nevertheless we have demonstrated that  $SO_2$ -photochemical perturbations have the potential to explain certain features of the tropical  $O_3$  changes observed after the Mount Pinatubo eruption. □

Received 11 November 1992; accepted 10 February 1993.

- Grant, W. B. *et al.* *Geophys. Res. Lett.* **19**, 1109–1112 (1992).
- Schoeberl, M. R., Bhartia, P. K., Hilsenrath, E. & Torres, O. *Geophys. Res. Lett.* **20**, 29–32 (1993).
- Hofmann, D. J. & Solomon, S. *J. geophys. Res.* **94**, 5029–5041 (1989).
- Prather, M. J. *J. geophys. Res.* **97**, 10187–10191 (1992).
- Michelangelo, D. V., Allen, M. & Yung, Y. L. *J. geophys. Res.* **94**, 18429–18443 (1989).
- Brasseur, G. P. & Granier, C. *Science* **257**, 1239–1242 (1992).
- Kinne, S., Toon, O. B. & Prather, M. J. *Geophys. Res. Lett.* **19**, 1927–1930 (1992).
- Chandra, S. *Geophys. Res. Lett.* **20**, 33–36 (1993).
- Bluth, G. J. S., Doiron, S. D., Schnetzler, C. C., Krueger, A. J. & Walter, L. S. *Geophys. Res. Lett.* **9**, 151–154 (1992).
- McCormick, M. P. & Veiga, R. E. *Geophys. Res. Lett.* **9**, 155–158 (1992).
- Stockwell, W. R. & Calvert, J. G. *Atmos. Environ.* **17**, 2231–2235 (1983).
- McKeen, S. A., Liu, S. C. & Kiang, C. S. *J. geophys. Res.* **89**, 4873–4881 (1984).
- Crutzen, P. J. & Schmailzl, U. *Planet. Space Sci.* **31**, 1009–1032 (1983).
- Okabe, H. *Photochemistry of Small Molecules* (Wiley-Interscience, New York, 1978).

- Harwood, R. S. & Pyle, J. A. *Q. J. R. met. Soc.* **101**, 723–748 (1975).
- Haigh, J. D. & Pyle, J. A. *Q. J. R. met. Soc.* **108**, 551–574 (1982).
- Bekki, S. & Pyle, J. A. *J. geophys. Res.* **97**, 15839–15847 (1992).
- DeMore, W. B. *et al.* *Chemical Kinetics and Photochemical Data for Use in Stratospheric Modeling* Evaluation 9, 90–1 (NASA/JPL Publ. 1990).
- Yung, Y. L. & DeMore, W. B. *Icarus* **51**, 199–247 (1982).
- Bernard, A., Demaiffe, D., Matelli, N., Runong-bayan, R. S. *Nature* **354**, 139–140 (1991).
- Trepte, C. R. & Hitchman, M. H. *Nature* **355**, 626–628 (1992).

ACKNOWLEDGEMENTS. We thank S. Chandra, M. R. Schoeberl and J. W. Waters for providing their data prior to publication. This work was supported by the UK Department of the Environment and the UK Universities Global Atmospheric Modelling (UGAMP) programme.

## Isotopic evidence for the source of lead in Greenland snows since the late 1960s

K. J. R. Rosman\*, W. Chisholm\*, C. F. Boutron†‡, J. P. Candelone† & U. Görlach†

\* Department of Applied Physics, Curtin University of Technology, Bentley 6102, Western Australia, Australia

† Laboratoire de Glaciologie et Géophysique de l'Environnement du CNRS, 54 rue Molière, Domaine Universitaire, BP 96, 38402 St Martin d'Hères Cedex, France

‡ UFR de Mécanique, Université Joseph Fourier de Grenoble, Domaine Universitaire, BP 69, 38401, Cedex, France

IN 1969, Murozumi *et al.*<sup>1</sup> demonstrated that the concentration of lead in Greenland snow had increased by a factor of 200 since ancient times, and concluded that most of this increase was a result of the use of alkyl-leaded petrol. Partly because of these findings, the United States and other western countries limited the use of lead additives in petrol from about 1970. Recently, Boutron *et al.*<sup>2</sup> showed that the lead concentration in Greenland snow had decreased by a factor of  $\sim 7.5$  over the past 20 years, and suggested that this was a result of the decline in use of leaded petrol. We present here measurements of the  $^{206}\text{Pb}/^{207}\text{Pb}$  ratio of the lead contained in the samples studied by Boutron *et al.* Because aerosols from the atmosphere above the United States are more radiogenic than those from Eurasia, we can trace the relative contributions of these two sources in the Greenland lead over the period analysed by Boutron *et al.* We find that the United States was a significant source of lead in the 1970s, but it has since declined considerably in relative importance. This decline mirrors the decrease in use of leaded petrol in the United States, confirming the earlier hypothesis.

We analysed aliquots of the samples collected by Boutron *et al.*<sup>2</sup>, who collected snow to a depth of 10.7 m at Summit ( $72^\circ 35' \text{N}$ ,  $37^\circ 38' \text{W}$ ) in central Greenland using a hand-operated auger, made of polycarbonate. Details of the elaborate low-contamination procedures used to drill, collect, process and store these samples are given in ref. 2. These samples were taken at an elevation of  $\sim 3.23$  km and give a more representative sample of tropospheric aerosols than those collected at sea level. This 10.7-m core provided a continuous record of snow deposited between 1967 and 1989, which includes the period when the consumption of leaded gasoline peaked in the Northern Hemisphere.

Lead has four naturally occurring stable isotopes, three of which are the end products of radioactive decay chains:  $^{238}\text{U} \rightarrow ^{206}\text{Pb}$ ,  $^{235}\text{U} \rightarrow ^{207}\text{Pb}$  and  $^{232}\text{Th} \rightarrow ^{208}\text{Pb}$ . Consequently, industrial lead will vary in isotopic composition depending on the primary source. In principle, it is therefore possible to identify the source of atmospheric lead if the isotopic history of the potential source regions is well documented.

Aliquots of decontaminated bulk samples<sup>2</sup> were transferred to acid-cleaned<sup>3</sup> bottles (made of low-density conventional polythene), frozen, then transported to Curtin University of Technology in Western Australia where they were analysed. The samples were processed inside a HEPA-filtered clean-air

TABLE 1 Estimates of lead concentrations in Greenland snow

|                                                                | 1968       | 1972       | 1976      | 1980      | 1984      | 1988      |
|----------------------------------------------------------------|------------|------------|-----------|-----------|-----------|-----------|
| $^{206}\text{Pb}/^{207}\text{Pb}$ ratios                       |            |            |           |           |           |           |
| Greenland snow*                                                | 1.159      | 1.18       | 1.182     | 1.182     | 1.164     | 1.155     |
| US aerosols†                                                   | 1.18       | 1.20       | 1.21      | 1.22      | 1.21      | 1.20      |
| Eurasia & Canada‡                                              | 1.14       | 1.14       | 1.14      | 1.14      | 1.14      | 1.14      |
| Concentration of Pb in Greenland snow ( $\mu\text{g g}^{-1}$ ) |            |            |           |           |           |           |
| Greenland (total)                                              | 135.5      | 118.4      | 94.2      | 63.3      | 55.3      | 33.7      |
| US                                                             | 64.2       | 78.8       | 56.8      | 32.9      | 19.1      | 8.4       |
| Eurasia & Canada                                               | 71.3       | 39.6       | 37.4      | 30.4      | 36.2      | 25.3      |
| Range‡                                                         | $\pm 16.0$ | $\pm 13.1$ | $\pm 8.1$ | $\pm 4.1$ | $\pm 2.7$ | $\pm 2.2$ |

Estimates are based on the measured isotope ratios and those in the principal contributing sources.

\* These ratios are a weighted average of the data shown in Fig. 1. Groups of points adjacent to the tabulated date have been averaged using  $R = \sum w_i R_i / \sum w_i$ , where  $w_i$  is the mass of Pb with ratio  $R_i$ .

† Estimated by averaging values given in cited literature (see text). The relatively small differences shown between the signatures of the two source regions and Greenland snow in 1968 makes the computed concentrations particularly sensitive to the ratios used.

‡ This range applies to both sources and was determined by allowing a change of  $\pm 0.01$  in the US isotopic ratio.

laboratory using ultra-clean conditions and procedures. Each sample was evaporated under a filtered stream of nitrogen in a PFA beaker together with  $\text{HNO}_3$ ,  $\text{H}_3\text{PO}_4$  and a  $^{205}\text{Pb}$  spike. By adding a  $^{205}\text{Pb}$  tracer it was possible to measure simultaneously the concentration and isotopic composition of the lead sample.

The sample was transferred together with  $\text{H}_3\text{PO}_4$  and silica gel to the rhenium filament of a thermal ion source, then analysed in a VG354 mass spectrometer using a Daly collector operating in analog mode. The average procedural blank for an isotopic analysis amounted to  $\sim 3\%$  of the sample analysed. A correction was made to the isotopic composition of the sample for this contribution, although the changes generally amounted to less than 0.3% of the  $^{206}\text{Pb}/^{207}\text{Pb}$  ratio. The accuracy of the mass spectrometer was monitored by analysing  $\sim 100$  pg amounts of the lead isotopic standard SRM981 at regular intervals. After making a correction of 0.31% per mass unit for isotopic fractionation, the  $^{206}\text{Pb}/^{207}\text{Pb}$  ratios agreed with the certified values to within  $\sim 0.4\%$  (95% confidence limits).

The samples were dated using the seasonal variations of methanesulphonic acid (M. Legrand, personal communication). The lead concentrations were measured by isotope dilution mass spectrometry to an accuracy of better than 20% (95% confidence limit). These concentrations were slightly lower than the values

given by Boutron *et al.*<sup>2</sup> on aliquots of the same bulk sample, but within errors, indicating that they were not contaminated by the storage containers or during the transfer process. To demonstrate that the coring device has not contaminated the sample, one normally measures the lead concentration profile of each core<sup>2,4</sup>. We measured one core and found no increase in concentration on the outside, but we did detect an increase of 0.5% in the  $^{206}\text{Pb}/^{207}\text{Pb}$  ratio. As the outside is the most contaminated part of the core, the effect on the inner samples can be expected to be negligible. We ruled out the possibility that lead contamination had occurred during subsampling by analysing ultra-pure water which had been allowed to stand in the subsampling tubes.

The results of measuring the isotopic composition of the lead in 25 snow core sections covering the last two decades are shown in Fig. 1 as a plot of the  $^{206}\text{Pb}/^{207}\text{Pb}$  ratio against time. There has been a general trend upward since 1967 reaching a maximum between 1975 and 1980, then a fall back to the 1967 values, and even lower, by 1986. The  $^{208}\text{Pb}/^{207}\text{Pb}$  and  $^{206}\text{Pb}/^{204}\text{Pb}$  ratios, which were also measured, were highly correlated with the  $^{206}\text{Pb}/^{207}\text{Pb}$  ratios and therefore showed a similar trend with time. A complete set of isotope ratios will be published elsewhere.

The United States, Eurasia and Canada are considered to be the main sources of polluted air reaching Greenland<sup>5</sup>. To assess the contribution from each of these sources, we need to know the isotopic signatures of the atmospheric aerosols from each of these areas. For US aerosols, the most comprehensive isotope data available between 1967 and 1974 (ref. 6) that the  $^{206}\text{Pb}/^{207}\text{Pb}$  ratio changed from 1.151 to 1.211 over this period. This change was attributed to the shift to the more radiogenic Mississippi Valley lead in alkyl-Pb additives used in US gasoline. Later values have remained fairly constant: 1.22 for 1980 (ref. 7); an average of 1.213 in 1982–84 (eastern US) and 1.221 in 1986 (US mid-west; both from ref. 8); 1.18–1.22 by 1989 (R. Flegal, personal communication) when the Summit snow core was taken. Canadian aerosols give a lower and fairly constant value: 1.165 in 1979 (in Vancouver<sup>9</sup>) and 1.148 in 1980 (in Victoria<sup>9</sup>); typical values of  $1.15 \pm 0.01$  between 1982 and 1986 (ref. 8). Church *et al.*<sup>10</sup>, however, report values of  $\sim 1.20$  by late 1988 in air originating in Canada, but collected over the Atlantic Ocean. Eurasian lead emissions are much greater than Canadian and the  $^{206}\text{Pb}/^{207}\text{Pb}$  ratio is poorly documented. Hopper *et al.*<sup>11</sup> have indicated a value of 1.156 for year 1988. Other data for the period 1965 to 1988 indicate an average value of 1.14 although there is wide scatter<sup>12–17</sup>. In summary, the isotopic signature of US lead changed from 1.15 in 1964 to 1.20 by 1974 and remained between 1.20 and 1.22 for the following two decades. Neither Eurasia nor Canada appears to have substantially altered its average signature (1.15 and 1.14, respectively) since the late 1960s, although Canadian lead may have risen to  $\sim 1.20$  after 1986.

On the basis of the above region signatures, a simple two-component linear model was constructed to derive the contributions of these components in the Greenland snow. US aerosol lead and a mixture of Eurasian and Canadian lead were used for the two isotopic signatures. A value of 1.14 was chosen for the latter. The results presented in Table 1 and Fig. 2 show that after 1972 there was a declining concentration in US lead and an almost constant concentration of the second (Eurasian/Canadian) component. Increasing the isotopic ratio of the second component to accommodate a significant proportion of Canadian lead does not alter the general shape of the concentration profiles. This establishes the US as a significant source of lead in the atmosphere during this period. This pattern also corresponds closely with that expected from leaded gasoline consumption figures (R. Fiat, personal communication).

The steep increase in the  $^{206}\text{Pb}/^{207}\text{Pb}$  ratio between 1967 and 1975 (Fig. 1) can be attributed to the influence of US lead emissions which were increasing both in quantity and isotopic

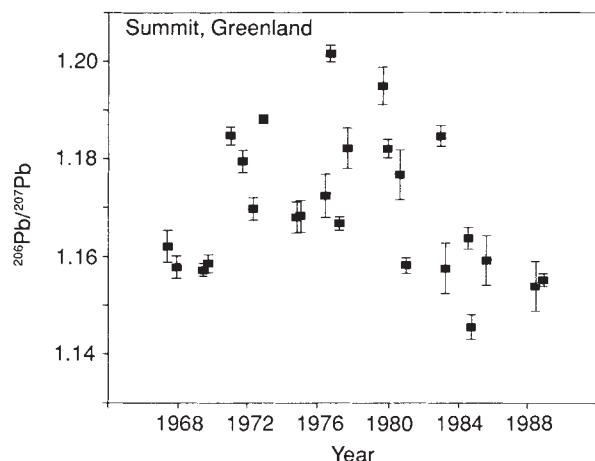


FIG. 1 Changes in the  $^{206}\text{Pb}/^{207}\text{Pb}$  ratio in Greenland snow between 1967 and 1988. The uncertainties given are 95% confidence intervals. The crescent-shaped trend is mainly due to combined changes in both the isotopic composition and the consumption of leaded gasoline in the United States.



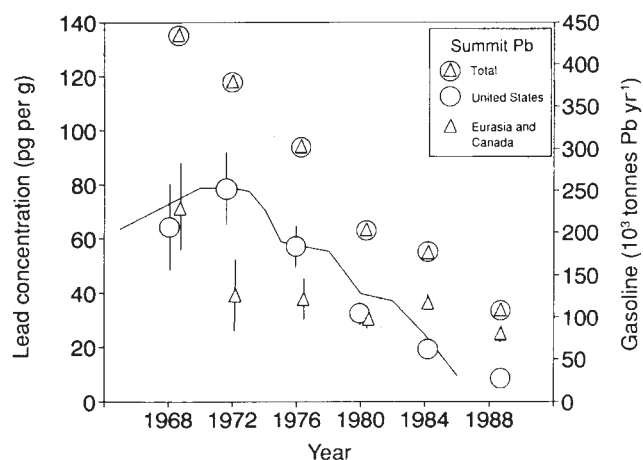


FIG. 2 Concentrations of the lead in Greenland snow (left-hand scale) from the principal source regions as a function of time. The data are taken from Table 1, and the uncertainty bars show the effect of altering the US  $^{206}\text{Pb}/^{207}\text{Pb}$  ratio by  $\pm 0.01$ . (The use of a higher  $^{206}\text{Pb}/^{207}\text{Pb}$  ratio for the Eurasian/Canadian component will lower the US and raise the Eurasian/Canadian concentrations, but the shapes of the two profiles do not change significantly.) The solid curve, showing US lead consumption in gasoline (right-hand scale), agrees well with the concentration profile at Summit.

ratio. The marked decrease after 1976 can be attributed to a decrease in the amount of US lead, whereas other major contributors in the Northern Hemisphere, such as Eurasia and Canada, maintained a lower but constant output.

Figure 1 also displays large variations in the  $^{206}\text{Pb}/^{207}\text{Pb}$  ratio between adjacent core sections which integrate 3–8 months of snow. These show that there are occasions when air masses from different source regions have fairly direct access to Summit. For example, two adjacent samples collected in 1976 had  $^{206}\text{Pb}/^{207}\text{Pb}$  ratios of 1.172 and 1.201, whereas two in 1979 had values of 1.185 and 1.158. The samples with the high ratios are clearly dominated by US lead, whereas the other samples contain a substantial component of less-radiogenic lead, probably from Eurasia or Canada. □

Received 6 November 1992; accepted 26 January 1993.

1. Murozumi, M., Chow, T. J. & Patterson, C. C. *Geochim. cosmochim. Acta* **33**, 1247–1294 (1969).
2. Boutron, C. F., Görlach, U., Candelone, J. P., Bolshov, M. A. & Delmas, R. J. *Nature* **353**, 153–156 (1991).
3. Boutron, C. F. *Fresenius Z. analyt. Chem.* **337**, 482–491 (1990).
4. Ng, A. & Patterson, C. *Geochim. cosmochim. Acta* **45**, 2109–2121 (1981).
5. Wolff, E. W. *Antarct. Sci.* **2**, 189–205 (1990).
6. Chow, T. J., Snyder, C. B. & Earl, J. L. *Proc. Conf. IAEA-191/4*, 95–108 (IAEA, Vienna, 1975).
7. Shirahata, H., Elias, R. W. & Patterson, C. C. *Geochim. cosmochim. Acta* **44**, 149–162 (1980).
8. Sturges, W. T. & Barrie, L. A. *Nature* **329**, 144–146 (1987).
9. Stukas, V. J. & Wong, C. S. *Science* **211**, 1424–1427 (1981).
10. Church, T. M. et al. *Global biogeochem. Cycles* **4**, 431–443 (1990).
11. Hopper, J. F., Ross, H. B., Sturges, W. T. & Barrie, L. A. *Tellus B* **43**, 45–60 (1991).
12. Hamilton, E. I. & Clifton, R. J. *Estuar. Coast. mar. Sci.* **8**, 271–278 (1979).
13. Maring, M., Settle, D. M., Buat-Ménard, P., Dulac, F. & Patterson, C. C. *Nature* **330**, 154–156 (1987).
14. Facchetti, S. et al. *Isotopic Lead Experiment* (Commission of the European Communities, US EPA 1982).
15. Petit, D., Mennessier, J. P. & Lamberts, L. *Atmos. Environ.* **18**, 1189–1193 (1984).
16. Elbaz-Poulichet, F., Hollinger, P., Huang, W. W. & Martin, J.-M. *Nature* **308**, 409–413 (1984).
17. Keinonen, M. *Sci. tot. Environ.* **113**, 251–268 (1992).

ACKNOWLEDGEMENTS. We thank C. Rado and R. Delmas for the field sampling and our colleagues in the Curtin University Isotope Science Group for discussions. This work was supported by the French Ministry of the Environment, the University of Grenoble, and the Commission of the European Communities and Switzerland as part of the EUROCORE program.

## Cambrian sea water preserved as inclusions in marine low-magnesium calcite cement

W. J. Johnson & R. H. Goldstein

Department of Geology, University of Kansas, Lawrence, Kansas 66045, USA

THE existence of temporal changes in the chemical composition of the oceans, which could provide constraints on the potential variability of the ocean–atmosphere system, remains an open question. Assessments of the chemistry of ancient oceans have relied largely on analysis of marine precipitates, generally carbonate and evaporite minerals<sup>1</sup>. These studies suggest that, whereas marine salinity has remained relatively stable over Phanerozoic time<sup>1</sup>, magnesium, calcium and sulphate concentrations of ancient oceans and  $\text{CO}_2$  partial pressure of ancient atmospheres may have changed<sup>2–4</sup>. The ratios of isotopes of carbon, oxygen, sulphur and strontium also appear to have varied<sup>3,5,6</sup>. Here we present analyses of primary, one-phase fluid inclusions in Cambrian and Ordovician marine cements, which appear to represent aliquots of early Palaeozoic oceans. The cements have trace element, stable isotope and strontium isotope contents that are consistent with their having been precipitated in a Cambrian–Ordovician marine environment, and the fluids have marine salinities. As these (apparently primary) cements are low-magnesium calcite, unlike the predominantly high-magnesium calcite and aragonite of today's carbonate precipitates, the chemistry of the Cambrian ocean–atmosphere system seems to have been different from that of today.

Recent studies have suggested that the dominant carbonate mineral precipitating from sea water may have varied through time<sup>7–10</sup>. Aragonite and high-magnesium calcite (HMC) are the dominant precipitates from modern shallow marine waters<sup>11</sup>. Petrographic evidence suggests, however, that calcites of

moderate to low magnesium content (LMC) may have been the dominant precipitate during some geological intervals<sup>8,12</sup>. The metastability of both aragonite and HMC, and their susceptibility to recrystallization to LMC, makes this hypothesis difficult to evaluate through petrographic methods. If such temporal variations exist, this would suggest that changes have occurred in one of the chemical parameters that control the type of carbonate mineral precipitating from sea water, such as  $\text{CO}_2$  partial pressure ( $p_{\text{CO}_2}$ ) or the Mg/Ca ratio<sup>13</sup>. The analysis of preserved inclusions of ancient sea water may allow such variations to be measured directly.

Here we look at three shallow marine truncation surfaces which bracket the Cambrian/Ordovician boundary, Wilberns Formation, Texas, USA. At these surfaces, calcite cements, ooids and fossil fragments are truncated and overlain by shallow marine sediment, indicating that the cements are syndepositional shallow marine precipitates and that these surfaces formed hard substrates on the sea floor (hardgrounds). The cements below these truncation surfaces are bladed (ratio of length to width is 1.5:1 to 6:1), isopachous fringes of low-magnesium calcite (0.5–2.0 mol %  $\text{MgCO}_3$ ). These cement crystals contain abundant one-phase liquid inclusions and less-common two-phase inclusions (Fig. 1). The existence of three-dimensional arrays of one-phase and two-phase inclusions that are confined by former growth surfaces of individual cement crystals indicates that these inclusions are primary<sup>14</sup>. One- and two-phase secondary inclusions are also present along rare healed fractures. Scanning electron microscope analysis did not reveal any microdolomite inclusions within the bladed cements.

Analysis of fluid inclusions within these Cambrian–Ordovician cements has focused on the one-phase primary inclusions. In general, primary inclusions which remain one-phase have survived thermal re-equilibration during burial heating and therefore retain their original fluid content, whereas associated two-phase inclusions have undergone re-equilibration<sup>15</sup>. Inclusion freezing point depressions ( $T_{\text{m,ice}}$ ) are shown

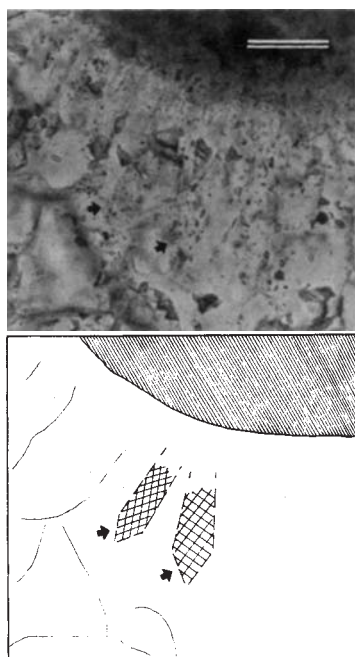


FIG. 1 Photomicrograph (top) of three-dimensional arrays of one-phase inclusions confined by former growth surfaces within individual cement crystals (hardground 2). The cross-hatched areas in the sketch below show the location of two such areas. Arrows are also provided for reference. Scale bar is 50  $\mu\text{m}$ .

in Fig. 2. Most of the values for one-phase inclusions lie between  $-1.7$  and  $-2.5$   $^{\circ}\text{C}$  (with a median of  $-2.1$   $^{\circ}\text{C}$ ); the mode is at  $-1.9$   $^{\circ}\text{C}$ . This corresponds to a salinity range of 31 to 47 parts per thousand (‰) sea water equivalent with the mode at 35‰ (the salinity of normal modern sea water)<sup>16</sup>. A few one-phase inclusions occur immediately to either side of the  $-1.7$  to  $-2.5$   $^{\circ}\text{C}$  range. These inclusions, however, contain fluids with salinities similar to those of all-liquid secondary inclusions present along fractures, and indicate either areas of minor recrystallization or refilling of primary inclusion cavities with secondary fluids. Rare one-phase inclusions with much lower  $T_{m,ice}$  values of  $-8.0$  to  $-14.0$   $^{\circ}\text{C}$  (12 to 18% NaCl equivalent) occur within the same salinity range as two-phase inclusions and almost certainly represent necking down of originally two-phase inclusions. This interpretation is supported by the presence of rare two-phase inclusions with significantly higher homogenization temperatures than the dominant mode (a few are in excess of 200  $^{\circ}\text{C}$ ).

The median freezing point depression for the primary one-phase inclusions corresponds to a salinity of 39‰, which is slightly above the normal marine mode in the data and is consistent with precipitation of some of these cements within a

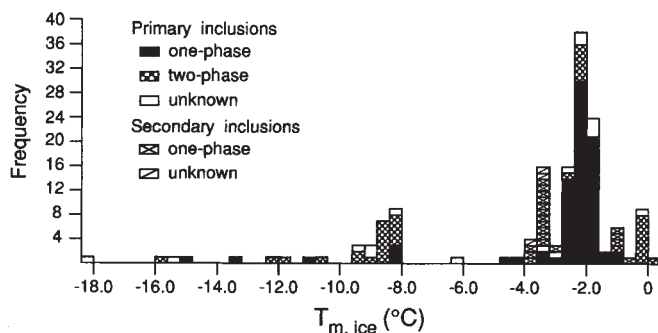


FIG. 2 Freezing point depression data from the three hardgrounds. Primary inclusion data are from bladed cements truncated by the hardground surface. Secondary inclusions are from a cement formed during burial. Fluid inclusion selection was biased towards one-phase inclusions. One-phase inclusions were stressed by heating to nucleate a vapour bubble before measurement. Data margin of error is better than 0.2  $^{\circ}\text{C}$ .

slightly restricted environment. Salinities in the modern Persian Gulf also lie in this range<sup>17</sup>. The observation that the primary fluid inclusions contain marine salinities, and have not been refilled with other fluids, is strong evidence that these cements are marine precipitates that have not undergone significant recrystallization.

Stable isotope, trace element and cathodoluminescence analyses of these cements allow us to test this further. The stable isotope data (Fig. 3) yield  $\delta^{18}\text{O}$  values that closely correspond to marine calcite values previously reported for the Cambrian and Ordovician periods<sup>18</sup>. The  $\delta^{13}\text{C}$  values are slightly more positive than data from an equivalent interval from Australia<sup>19</sup>, but all values are within the range of variation seen in modern shallow marine calcites<sup>20</sup>.

Trace element analyses of the cements were made with an electron microprobe. Strontium concentrations ranged from 207 to 2,201 p.p.m. (three analyses below the minimum detection limit (MDL) of 195 p.p.m.; mean above MDL was 597 p.p.m.). Manganese concentrations were below the MDL of 273 p.p.m. in hardgrounds 2 and 3 and ranged from below MDL to

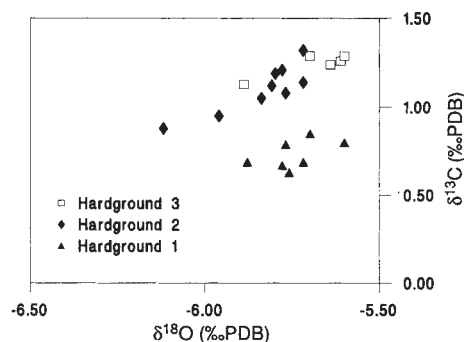


FIG. 3 Stable isotope data for hardground cements. Hardground 1 is Late Cambrian in age, hardground 2 is at the Cambrian/Ordovician boundary, and hardground 3 is Early Ordovician in age. Microsampled cements were analysed for stable isotopes using standard mass spectrometric techniques.  $\delta^{13}\text{C} = [(^{13}\text{C}/^{12}\text{C})_{\text{sample}} / (^{13}\text{C}/^{12}\text{C})_{\text{PDB}}] - 1$  and  $\delta^{18}\text{O} = [(^{18}\text{O}/^{16}\text{O})_{\text{sample}} / (^{18}\text{O}/^{16}\text{O})_{\text{PDB}}] - 1$ .

1,840 p.p.m. in hardground 1 (see Fig. 3 for hardground positions). Iron content ranged from below detection (MDL = 216) to 872 p.p.m.. Strontium concentrations are comparable to what would be expected for a marine precipitate. The broad range in strontium values is comparable to the range observed in modern marine high-magnesium calcites<sup>21</sup>, but contrasts markedly with the small spread in magnesium values. This does not conform to models correlating strontium and magnesium uptake during marine calcite precipitation<sup>22</sup>. To explain the Cambrian–Ordovician cement data by invoking recrystallization, however, requires a diagenetic process relatively closed with respect to strontium and the stable isotopes of oxygen and carbon. Such a process would require the sequestering of magnesium, probably in the form of microdolomite inclusions<sup>23</sup>. The absence of microdolomite inclusions in these cements argues against this mechanism. Moreover, preservation of fluid inclusions with marine salinities precludes open-system recrystallization in non-marine fluids. An alternative model involving open-system recrystallization in deep (cold) marine waters<sup>21,24</sup> is also largely excluded by the shallow marine setting and extended range of salinities (31 to 47‰) exhibited by these one-phase inclusions.

Iron concentrations are higher than would be expected of a precipitate from fully oxygenated sea water, but this does not necessarily imply recrystallization. High iron contents may, instead, indicate cement precipitation in suboxic microenviron-



ments with localized iron sources within the sediment<sup>21,25</sup>. This view is supported by the lack of a negative correlation between strontium and iron values, which could be predicted if recrystallization were prevalent<sup>26</sup>.

Manganese concentrations increase progressively towards pore centres in hardground 1. This trend is associated with a change in the cathodoluminescence emission of bladed cements from nonluminescent (with some very dull luminescence) to brightly luminescent. Inclusions with seawater salinities exist in both nonluminescent and brightly luminescent crystals, indicating that the manganese trend toward pore centres may also be a result of seawater precipitation in oxic to suboxic microenvironments (probably as a result of progressive pore restriction combined with continuing decay of organic carbon). This hypothesis is supported by stable isotope data which show  $\delta^{18}\text{O}$  values similar to other hardground cements but  $\delta^{13}\text{C}$  values  $\sim 0.4\%$  more negative. Bladed cements within hardgrounds 2 and 3 show very dull luminescence or none at all.

We measured  $^{87}\text{Sr}/^{86}\text{Sr}$  in bladed cements from hardgrounds 1 and 3, normalizing them to a value of 0.71014 for NBS 987. Normalized values were 0.708958 for hardground 3 and 0.708970 for hardground 1 ( $\pm 0.00003$  at the 95% confidence level). These are very close to those previously estimated for the Late Cambrian to Early Ordovician interval<sup>6</sup>. Again, this is evidence that the cements have not recrystallized appreciably (within an open system) and that they, and their contained fluid inclusions, preserve primary marine signatures from the early Palaeozoic.

Petrographic relationships indicating early (syndepositional) erosional truncation of these cements, with the preservation of inclusion-rich areas that are confined by former growth surfaces, provide strong evidence that they were initially precipitated during the Cambrian–Ordovician. Our analyses yield no definitive evidence of significant recrystallization and in fact provide evidence against it. No single geochemical signature can be considered 'definitive', but in combination they provide strong evidence that the cements and their fluid inclusions are relatively pristine. We conclude that the cements formed as low-magnesium calcite marine precipitates at (and below) the sediment–water interface and have not undergone significant recrystallization.

Our results support the hypothesis that the dominant mineral precipitating from sea water has changed through time<sup>7–10</sup>, which implies that some chemical parameter in the atmosphere–seawater system (significant with respect to calcite precipitation) has changed between Cambrian time and the modern. The partial pressure of  $\text{CO}_2$  may be one such parameter, with modelling studies suggesting that it was much higher during the Cambrian and Ordovician than at any other time in the Phanerozoic<sup>4</sup>. Changes in the Mg/Ca ratio of sea water have been proposed as an alternative mechanism<sup>18,27</sup>. Samples of ancient sea water preserved as fluid inclusions could provide a direct means of answering these and other questions. □

20. Gonzalez, L. A. & Lohmann, K. C. *Geology* **13**, 811–814 (1985).
21. Major, R. P. & Wilber, R. J. *Bull. geol. Soc. Am.* **103**, 461–471 (1991).
22. Carpenter, S. J. & Lohmann, K. C. *Geochim. cosmochim. Acta* **56**, 1837–1849 (1992).
23. Lohmann, K. C. & Meyers, W. J. *J. sedim. Petrol.* **48**, 475–488 (1977).
24. Schlager, W. & James, N. P. *Sedimentology* **25**, 675–702 (1978).
25. Weiss, C. P. & Wilkinson, B. H. *J. sedim. Petrol.* **58**, 468–478 (1988).
26. Brand, U. & Veizer, J. *J. sedim. Petrol.* **50**, 1219–1236 (1980).
27. Wilkinson, B. H. *Geology* **7**, 524–527 (1979).

ACKNOWLEDGEMENTS. We thank J. Miller, Mr and Mrs D Lange, and Mr and Mrs H. Weige for their help. This work was partially supported by KU GRF, GSA and NSF.

## Holocene dwarf mammoths from Wrangel Island in the Siberian Arctic

S. L. Vartanyan\*, V. E. Garutt† & A. V. Sheret‡||

\* Wrangel Island State Reserve, 686870 Ushakovskoye, Magadan Region, Russia

† Zoological Institute, Russian Academy of Sciences, 1 Universitetskaya naberezhnaya, 199034 St Petersburg, Russia

‡ Severtsov Institute of Evolutionary Animal Morphology and Ecology, Russian Academy of Sciences, 33 Leninskiy Prospekt, 117071 Moscow, Russia

THE cause of extinction of the woolly mammoth, *Mammuthus primigenius* (Blumenbach), is still debated. A major environmental change at the Pleistocene–Holocene boundary, hunting by early man, or both together are among the main explanations that have been suggested. But hardly anyone has doubted that mammoths had become extinct everywhere by around 9,500 years before present (BP). We report here new discoveries on Wrangel Island in the Arctic Ocean that force this view to be revised. Along with normal-sized mammoth fossils dating to the end of the Pleistocene, numerous teeth of dwarf mammoth dated 7,000–4,000 yr BP have been found there. The island is thought to have become separated from the mainland by 12,000 yr BP. Survival of a mammoth population may be explained by local topography and climatic features, which permitted relictual preservation of communities of steppe plants. We interpret the dwarfing of the Wrangel mammoths as a result of the insularity effect, combined with a response to the general trend towards unfavourable environment in the Holocene.

The woolly mammoth was well adapted to the rigorous climate of the Ice Age. As recently as 15,000–20,000 yr ago it inhabited northern and temperate latitudes of the Northern Hemisphere. At the end of the Pleistocene, about 10,000–12,000 yr ago, it disappeared from all the three continents almost simultaneously. No reliable radiocarbon ( $^{14}\text{C}$ ) dates younger than 12,000 yr BP obtained directly on woolly mammoth fossils are known from Europe or North America<sup>1–6</sup>, but there are 11  $^{14}\text{C}$  dates within the range 12,000 to 9,600 yr BP from the Siberian Arctic (A.V.S. and L. D. Sulerzhitzky, manuscript in preparation). On Wrangel Island (Fig. 1), an abundance of mammoth fossils was reported long ago<sup>7,8</sup>, but they have only been specifically collected since 1989 (by S.L.V.). Mammoth fossils were mainly picked from the surface and at a small depth in slope or valley-bottom sediments; some were found just in river channels. The first set of  $^{14}\text{C}$  dates from four tusk samples and one tibia bone (produced by the Laboratory of the Geographic Research Institute, St Petersburg (Leningrad) State University (LGU)) fall within the range 7,380 to 4,740 yr BP<sup>9</sup>. Later, L. D. Sulerzhitzky (Geological Institute, Russian Academy of Sciences, Moscow (GIN)) dated a tusk fragment, previously brought from Wrangel, to 5,980  $\pm$  90 yr BP (GIN-6310).

The LGU and GIN laboratories have now produced a total of 30 dates on mammoth tusks and bone samples collected by

Received 31 August 1992; accepted 5 February 1993.

1. Holland, H. D. *The Chemical Evolution of the Atmosphere and Oceans* (Princeton Univ. Press, 1984).
2. Wilkinson, B. H. & Algeo, T. J. *Am. J. Sci.* **239**, 1158–1194 (1989).
3. Clapprol, G. E., Holser, W. T., Kaplan, I. R., Sakai, H. & Zak, I. *Chem. Geol.* **28**, 199–260 (1980).
4. Berner, R. A. *Science* **249**, 1382–1386 (1990).
5. Lohmann, K. C. & Walker, J. C. G. *Geophys. Res. Lett.* **16**, 319–322 (1989).
6. Burke, W. H. *et al.* *Geology* **10**, 516–519 (1982).
7. Sandberg, P. A. *Nature* **305**, 19–22 (1983).
8. Mackenzie, F. T. & Pigott, J. D. *J. geol. Soc. Lond.* **138**, 183–196 (1981).
9. Wilkinson, B. H., Owen, R. M. & Carroll, A. R. *J. sedim. Petrol.* **55**, 171–183 (1985).
10. Palmer, T. J., Hudson, J. D. & Wilson, M. A. *Nature* **335**, 809–810 (1988).
11. Bathurst, R. G. C. *Carbonate Sediments and their Diagenesis* (Elsevier, Amsterdam, 1975).
12. Wilkinson, B. H. *J. geol. Educ.* **30**, 189–203 (1982).
13. Sandberg, P. A. in *The Carbon Cycle and Atmospheric  $\text{CO}_2$ : Natural Variations Archaean to Present* (eds Sundquist, E. T. & Broecker, W. S.) 585–594 (Am. Geophys. Un., Washington DC, 1985).
14. Roedder, E. *Fluid Inclusions* (Min. Soc. of America, Washington DC, 1984).
15. Goldstein, R. H. *J. sedim. Petrol.* **60**, 282–292 (1990).
16. Lyman, J. & Fleming, R. H. *J. mar. Res.* **3**, 134–146 (1940).
17. Purser, B. H. & Seibold, E. in *The Persian Gulf: Holocene Carbonate Sedimentation and Diagenesis in a Shallow Epicontinental Sea* (ed. Purser, B. H.) 1–9 (Springer, Berlin, 1973).
18. Carpenter, S. J. *et al.* *Geochim. cosmochim. Acta* **55**, 1991–2010 (1991).
19. Ripperdan, R. L., Margaritz, M., Nicoll, R. S. & Shergold, J. H. *Geology* **20**, 1039–1042 (1992).

|| To whom correspondence should be addressed.

S.L.V. on Wrangel. Twenty-nine are of Holocene age (3,730–7,620 yr BP) and the remaining sample, 12,750 yr BP (S.L.V., L. D. Sulerzhitsky and T. V. Tertychnaya, manuscript in preparation).

In 1991, S.L.V. collected a series of mammoth cheek teeth which were then studied by V.E.G. and A.V.S. The teeth are moderately preserved, about half are virtually undamaged, and many just slightly abraded. The whole sample (29 teeth and fragments) falls into two size classes. One comprises 5 specimens

of relatively large size, the other 24 of very small size, though belonging to adults (Fig. 2; Table 1a).

The larger teeth from Wrangel are the size of normal Late Quaternary mammoths from north-east Siberia, including the neotype of *Mammuthus primigenius*, the Taimyr mammoth skeleton<sup>10</sup>. The small teeth from Wrangel are about 20–25% smaller. The hypothesis suggested by this important size difference was that the small teeth from Wrangel belonged to the Holocene survivors previously dated by the tusks. This was

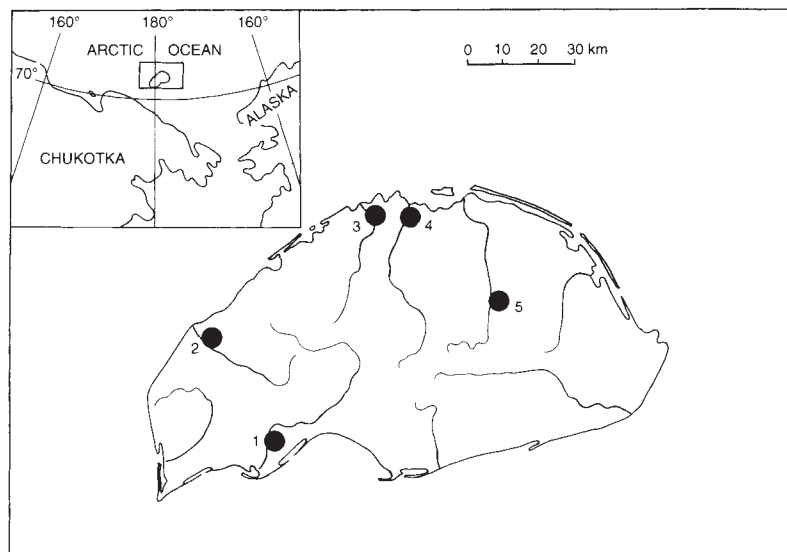
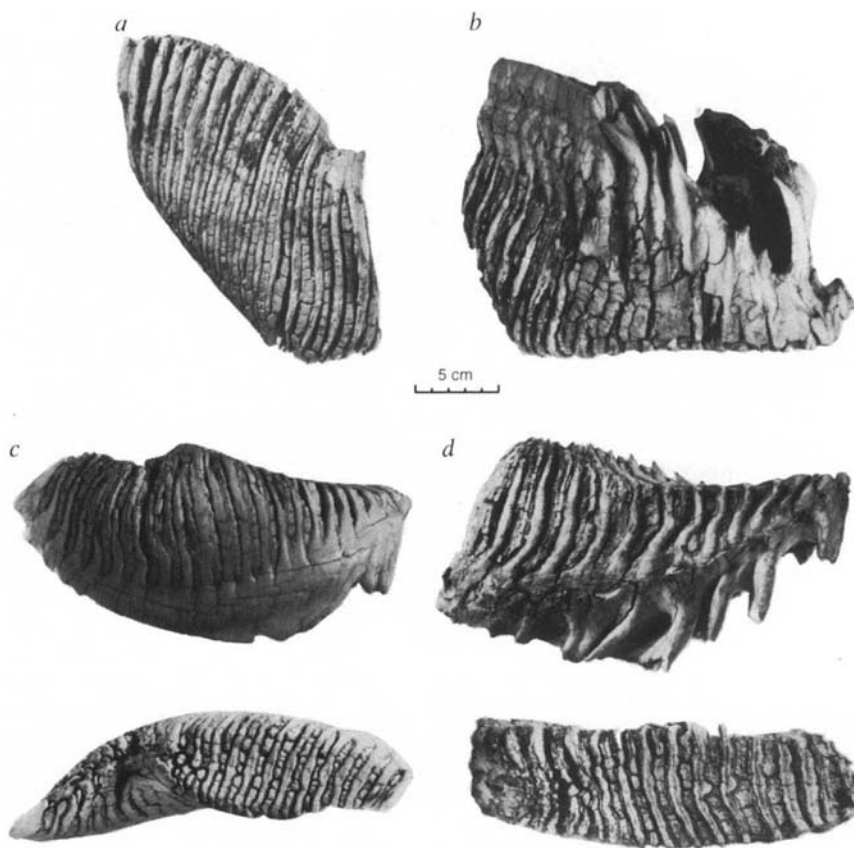


FIG. 1 Wrangel Island and the main areas of collecting mammoth teeth. (1) Mamontovaya River; (2) Goosinaya River; (3) Tundrovaya River; (4) Neizvestnaya River; (5) Krasnyy Flag River. The island is separated from the mainland, Chukotka, by Long's Strait with a minimum width of 150 km and maximum depth of 50 m. The central part of the island consists of low mountains (up to 1,000 m). Plateaus elevated up to 200–400 m lie west and east of the mountains; the lowland Tundra of Academy (up to 60 m) occupies the north of the island. Quaternary deposits of Wrangel Island are rather shallow and poorly studied. Marine sediments of uncertain age make up the northern lowland. Gravels, sands and silts, deposited by fluvial and slope processes, cover the plain surfaces and fill valley bottoms. Sometimes they include thick peat layers<sup>26</sup>. There are no traces of extensive glaciation of the island (S.L.V. manuscript in preparation).

FIG. 2 Last molar teeth ( $M_3$ ) of two forms of mammoth from Wrangel Island. *a*, Upper  $M_3$  of dwarf form, no. MAM-4, Mamontovaya River, in lateral view. This is one of four upper  $M_3$  teeth of the dwarf form in the collection. The tooth was just coming into wear when the animal died. Similar dental stage in living elephants (eruption of the last generation of molars) suggests an age of not less than about 30 yr<sup>24,27</sup>. Two more upper teeth (not illustrated) belong to the last but one generation ( $M_2$ ); a further one is either  $M_2$  or  $M_3$ . The pictured specimen, MAM-4, is the best preserved one, though several plates at the posterior end are missing (18 are preserved). The estimated total length of the crown was hardly more than 220 mm. Three teeth of similar size from Mamontovaya River have <sup>14</sup>C ages of 4,000–5,000 yr BP. *b*, Upper  $M_3$  of large form, no. KRF-1, Krasnyy Flag River, in lateral view. The tooth is deeply worn: the anterior roots have been resorbed and the corresponding part of the crown is lost. Only 20 plates are preserved, but their original number is estimated to have been 26 or 27 (ref. 28). The original crown length was more than 300 mm (268 mm preserved). The left  $M_3$  of the same individual (KRF-2) is dated as  $12,980 \pm 80$  yr BP. *c*, Lower  $M_3$  of dwarf form, no. GUS-10, Goosinaya River, in medial (image here reversed to facilitate comparison) and occlusal views. There are 15 measurable lower teeth of dwarf form in Wrangel collection; 12 of them, including GUS-10, certainly belong to the last generation (that is  $M_3$ ). The tooth GUS-10 is the best preserved one and has the maximum number (23) of preserved plates. Two anterior plates seem to be absent due to rather deep wear. All the other  $M_3$ s of the dwarf form are much more deeply worn, and belong to old or senescent individuals. Because of this their crown length and total number of plates cannot be reconstructed. *d*, Lower  $M_3$  of large form, no. KRF-3,



Krasnyy Flag River, in medial and occlusal views. The large tooth KRF-3 is also quite worn and lacks at least 10 plates (14 preserved).



TABLE 1 Results of analysis of mammoth teeth

(a) Maximum crown width of last molars of mammoth (mm)

|                                 | Upper M <sup>3</sup> |      |    | Lower M <sub>3</sub> |      |    |
|---------------------------------|----------------------|------|----|----------------------|------|----|
|                                 | Range                | Mean | n  | Range                | Mean | n  |
| Wrangel, small                  | 63–72                | 68.0 | 4  | 54–74                | 64.6 | 12 |
| Berelyokh, small                | 64–80                | 74.6 | 10 | 72–85                | 78.8 | 6  |
| Berelyokh,<br>whole sample      | 64–90                | 77.2 | 12 | 72–96                | 83.2 | 9  |
| Other NE Siberian<br>localities | 66–102               | 88.7 | 18 | 71–96                | 82.8 | 14 |
| Wrangel, large                  | 88–94                | 91.3 | 3  | 88;88                | 88.0 | 2  |
| Neotype <i>M.p.</i> , Taimyr    | 99;99                |      |    | 87;91                |      |    |

(b) Radiocarbon dates of mammoth teeth from Wrangel island

| Small form: |              |                 |                     |       |
|-------------|--------------|-----------------|---------------------|-------|
| LU-2798     | 4,010 ± 50   | Mamontovaya R.  | last upper molar    | MAM-6 |
| LU-2808     | 4,040 ± 30   | Mamontovaya R.  | tooth fragment      | MAM-2 |
| LU-2794     | 5,110 ± 40   | Mamontovaya R.  | last lower molar    | MAM-5 |
| LU-2799     | 6,260 ± 50   | Goosinaya R.    | last lower molar    | GUS-9 |
| LU-2810     | 6,890 ± 50   | Goosinaya R.    | tooth fragment      | GUS-7 |
| LU-2809     | 7,250 ± 60   | Tundrovaya R.   | last(?) lower molar | PIK-1 |
| Large form: |              |                 |                     |       |
| LU-2792     | 12,980 ± 80  | Krasnyy Flag R. | last upper molar    | KRF-2 |
| LU-2807     | 20,000 ± 110 | Neizvestnaya R. | last upper molar    | NZV-1 |

a. Crown width was used as the main size parameter because of either deep wear or fragmentation of most specimens. Berelyokh 'mammoth cemetery' in the lower Indigirka is one of the latest large samples of Siberian mammoth (12,850 ± 110–12,000 ± 130 yr BP). Most mammoth last molars from the Berelyokh collection are relatively small (particularly evident in small crown width<sup>23</sup>), but some are larger, close to normal late north-east Siberian mammoth. It is debated whether all the small-sized mammoth fossils from Berelyokh belong to females, or some are dwarfed males (see ref. 24 for references). The latter hypothesis may suggest that the large fossils are of earlier geologic age. That is why we present the whole Berelyokh sample and the small-sized teeth separately. Whether the small teeth from Berelyokh represent females only, or a whole dwarfed population, the Wrangel teeth are notably smaller in size even than them. b. The datings were made in the LGU laboratory on collagen by standard procedures<sup>25</sup>. As far as most samples were collected on the surface, the issue of possible contamination by younger carbon was carefully studied. Extensive methodological research entailing sequential assays (L. Sulerzhitsky, personal communication) has demonstrated that the thorough procedures of collagen extraction and cleaning removed all extraneous carbon from the samples before measurement. The validity of the Wrangel dates is further confirmed by comparison with more than 50 mainland mammoth fossils from Taimyr peninsula (A.V.S. and L. D. Sulerzhitsky, manuscript in preparation). Although these were also surface finds, none gave a date younger than 9,500 yr BP, and some proved to be older than 50,000 yr BP. Only the small Wrangel remains, of analogous provenance and preservation, and using the same techniques, gave dates younger in the range 7,000–4,000 yr BP. We are therefore confident in ruling out the great contamination which would be required spuriously to produce such young dates.

confirmed by the <sup>14</sup>C dating of six teeth and measurable fragments of the small form, and two of the large (Table 1b).

Because all dated small teeth are Holocene, and both large ones are late Pleistocene, there is no reason to suggest that they belong to males and females in the same population. Small last molars (M3) from Wrangel are among the smallest for *Mammuthus* (Fig. 3). Though we have no evidence on the size of limb bones, the small mammoth from Wrangel can be considered as a dwarfed form.

The existence of dwarfed forms of Pleistocene elephants on islands is well known<sup>11</sup>. Pigmy forms of straight-tusked elephant, *Palaeoloxodon*, inhabited some islands in the Mediterranean<sup>12,13</sup>. *Mammuthus exilis*, a dwarfed descendant of the North American mammoth *M. columbi*, existed on Santa-Rosa and other islands off the Californian coast at the end of the Pleistocene<sup>14,15</sup>. The existence of dwarfed woolly mammoth in northern Siberia at the end of the Pleistocene or even in early Holocene had been suggested by other workers<sup>16</sup>, but could not be proved.

It is not possible to reconstruct the body size of the Wrangel small form based on teeth alone. Note, however, that these teeth are about 30% smaller than those of the type *M. primigenius* from Taimyr, which was rather small itself (2.65 m shoulder height<sup>17</sup>). In many dwarfed forms body size was reduced to a much greater extent than the teeth<sup>18</sup>. This suggests that we can consider the 30% tooth reduction as only a minimum estimate for body size, and that the Holocene Wrangel mammoth was

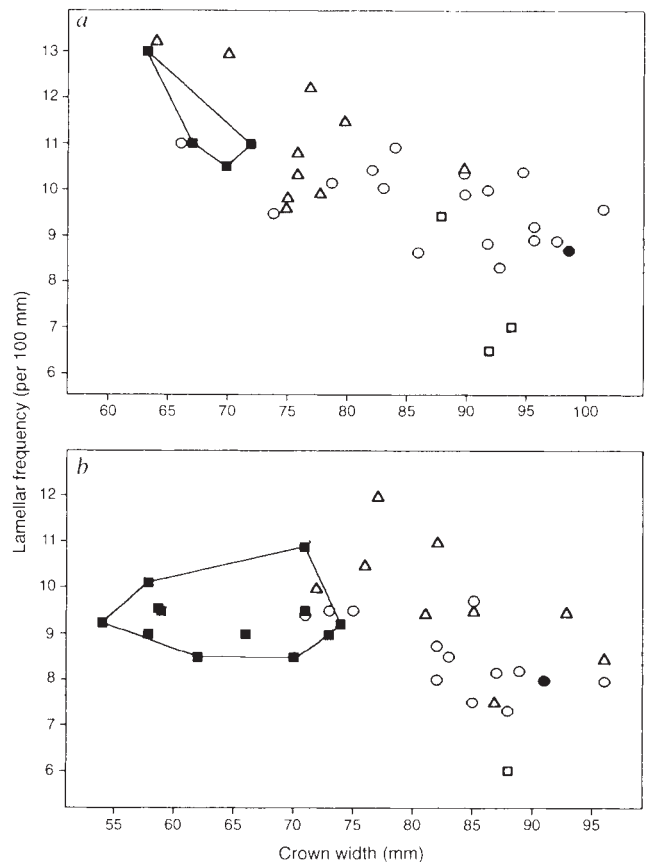


FIG. 3 Comparison of small and large forms of mammoth from Wrangel Island with some mainland samples. Maximum crown width is plotted against lamellar frequency (average number of plates in 100 mm of crown length). a, Last upper molars M<sup>3</sup>; b, last lower molars M<sub>3</sub>. Wrangel Island, small form (■); Wrangel Island, large form (□); Berelyokh 'Mammoth Cemetery' (△); other late Pleistocene localities of the Indigirka-Kolyma Lowland from A.V.S. collection (○); Taimyr Mammoth, neotype of *Mammuthus primigenius*<sup>10</sup> (●). Original data by V.E.G. and A.V.S. Measuring methods from ref. 29. Upper teeth of small Wrangel form display a very high lamellar frequency, like those from Berelyokh, which can be related to their small size and resulting lamellar compression<sup>30</sup>. Relatively low values of this index in lower teeth of the Wrangel dwarf are related to their deep wear (in lower teeth the plates diverge to the base of the crown, resulting in fewer plates per 100 mm on heavily worn teeth, that is, in lower lamellar frequencies).

very small.

During the late Pleistocene, until at least 16,000 yr BP, Wrangel Island together with the coastal lowlands of Arctic East Siberia, Alaska and what is now the vast and shallow Arctic shelf, was a part of the extensive land of Beringida<sup>31</sup>. According to reconstructions based on age estimation of submerged shorelines in the Bering Strait area<sup>19</sup>, Wrangel Island could still have been joined with the mainland as late as 13,000 yr BP, but by 12,000 yr BP that connection had been broken. The normal size of Wrangel mammoth teeth dating from the Pleistocene agrees with these estimates: it seems that the local population was not isolated from the main range until 12,000 yr BP. Later there is a gap in the mammoth record on the island until 7,000 yr BP, when the dwarfing had already occurred. This seems to have been enough time for insular dwarfing, comparable to the dwarfing of red deer on the island of Jersey (UK), which took less than 6,000 years<sup>18</sup>. But the trend of general environmental change must also be considered. As early as about 12,000 yr BP, that trend resulted in the disappearance of mammoth over most of its vast range, and survival only in the High Arctic, north of 70° N (A.V.S. and L.D. Sulerzhitsky, manuscript in preparation and ref. 20). At that time at least some Arctic mainland populations, such as that of the Berelyokh in the lower course of the

Indigirka River, show a reduction in body size (Table 1a). The period 9,500–8,000 yr BP is considered as the Holocene 'optimum' for this region, marked by increased temperature and humidity, and related turnover in plant communities<sup>20</sup>. By 9,500 yr BP, mammoth had vanished totally from the mainland and most of the Arctic islands, and it is likely that this time was also critical for the Wrangel Island mammoth.

How was a single mammoth population able to survive such hard times and why just on Wrangel Island? A possible explanation comes from present-day botanical observations<sup>21,22</sup>. Unlike the other islands within the Arctic Tundra subzone, Wrangel retains a much higher diversity of herb species, steppe plants in particular. Grasses, sedges and forbs, mainly xerophytes and cryoxerophytes in their ecology, play an important role in cryophyte steppe and tundra meadows. These herbaceous communities occupy some favourable habitats on the island, mostly in the basins and plateaus of its mountainous part, supported by more continental climate and abundance of carbonate rocks, and are considered as depauperated relicts of the Pleistocene grassland (tundra-steppe)<sup>22</sup>. It is likely that the island provided a relict of tundra-steppe habitat in the early Holocene too, but that is still to be proved.

Although oppressed and dwarfed, woolly mammoth could exist on Wrangel Island until at least 3,700 years ago. There is no evidence that it was hunted by man. The extraordinary fact is not that mammoth eventually became extinct on Wrangel like elsewhere, but that, together with some relicts of tundra-steppe flora, it was able to survive the early Holocene environmental revolution in this Arctic island refuge. □

## A Lyme borreliosis cycle in seabirds and *Ixodes uriae* ticks

Björn Olsén\*†, Thomas G. T. Jaenson‡, Laila Noppa\*, Jonas Bunikis\* & Sven Bergström\*§

\* Department of Microbiology, and † Department of Infectious Diseases, University of Umeå, S-90187 Umeå, Sweden

‡ Department of Zoology, University of Uppsala, Box 561, S-75122 Uppsala, Sweden

§ To whom correspondence should be addressed

THE Lyme disease spirochaete, *Borrelia burgdorferi* s.l., is the only *Borrelia* known to infect both mammals and birds<sup>1</sup>. The main vertebrate reservoirs of *B. burgdorferi* are thought to be various small and intermediate size mammals<sup>2</sup>, but the importance of birds as a reservoir has not been thoroughly explored. In the Northern and Southern Hemispheres the seabird tick, *Ixodes uriae*, is prevalent and closely associated with many species of colony-nesting marine birds<sup>3</sup>. Here we report the presence of spirochaetes, demonstrated by immunofluorescent assay, by polymerase chain reaction and in culture, in *I. uriae* infesting razorbills on an island in the Baltic Sea. This island is free from mammals. The protein profile of the spirochaetes and the sequences of their flagellin and *ospA* genes are identical to those of the Lyme disease spirochaete, *Borrelia burgdorferi* s.l., previously isolated from *I. ricinus* on a nearby island. In biopsies from the foot web of razorbills, *B. burgdorferi*-specific DNA was detected after amplification by polymerase chain reaction. Our results suggest that birds play an important part in the maintenance of *B. burgdorferi* and that mammals may not be a prerequisite for its life cycle.

There are several seabird colonies in the Baltic Sea and one of the largest is on the island of Bonden (63° 26' N, 20° 03' E; surface area, 0.45 km<sup>2</sup>) 12 km from the Swedish mainland. During the breeding season this island holds about 4,000 pairs of razorbill (*Alca torda*), 800 pairs of guillemot (*Uria aalge*), 500 pairs of black guillemot (*Cephus grylle*), and 50 pairs of gulls (*Larus* spp)<sup>4</sup>. The vegetation is dominated by herbs and grasses. Bonden is not inhabited by terrestrial mammals<sup>5</sup>.

During bird ringing on Bonden in July in 1991 and 1992, 145 razorbills, 223 guillemots, two black guillemot nestlings and five herring gull (*Larus argentatus*) fledglings were checked for ticks. Twenty-three adult female and fourteen nymphal fed *I. uriae* were collected from the feet of 13 razorbills and 2 guillemot

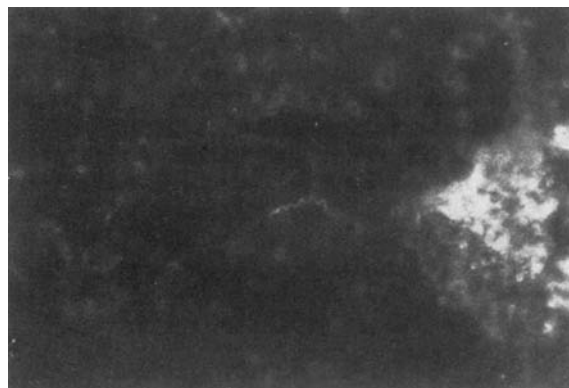


FIG. 1 *Borreliae* in mid-gut tissues of *Ixodes uriae*. After immersion in 70% ethyl alcohol, each tick was washed in sterile water. The content of the tick's idiosoma was excised aseptically into a few drops of sterile phosphate-buffered saline on a glass slide. Smears remaining on the slides after microscopy were dried and fixed and stained by indirect immunofluorescence using the anti-flagellum monoclonal antibody H9724 (ref. 8). After incubation for 30 min at 37 °C, slides were washed and incubated with goat anti-mouse (Becton Dickinson, UK) fluorescein isothiocyanate-labelled conjugate.

Received 15 October 1992; accepted 22 January 1993.

1. Aaris-Sørensen, K., Petersen, K. S. & Tauber, H. *Geol. Surv. Denmark* **14**, 1–44 (1990).
2. Lister, A. M. *CBA Res. Rep.* **77**, 51–59 (1991).
3. Goulet, J. A. J., Hedges, R. E. M., Law, I. A. & Perry, C. *Archaeometry* **29**, 125–155 (1987).
4. Lavrov, A. V. & Mashchenko, E. N. *Priroda* **1**, 52–55 (1991).
5. Kurtén, B. & Anderson, E. *Pleistocene Mammals of North America* (Columbia Univ. Press, New York, 1980).
6. Guthrie, R. D. *Frozen Fauna of the Mammoth Steppe: The Story of Blue Babe* (University of Chicago Press, Chicago, 1990).
7. Gekker, R. F. *Vest. Akad. Nauk SSSR* (in Russian) **1**, 60–62 (1938).
8. Markov, K. K. in *Extreme North-East of the U.S.S.R. Vol. 1 Wrangel Island* (in Russian) 25–48 (USSR Acad. Sci. Publishing, 1952).
9. Vartanyan, S. L., Arslanov, Kh. A., Tertychnaya, T. V. & Chernov, S. B. in *Secular Dynamics of Biogeocenoses. The 10th Sukachev Readings* (in Russian) 52–53 (Nauka, Moscow, 1992).
10. Garutt, W. E., Gentry, A. & Lister, A. M. *Bull. zool. Nomencl.* **47**, 38–44 (1990).
11. Sondaar, P. Y. in *Major Patterns in Vertebrate Evolution* (eds Hecht, M. K., Goody, P. C. & Hecht, B. M.) 671–706 (Plenum, New York, 1977).
12. Bate, D. M. A. *Proc. R. Soc.* **73**, 498–500 (1903).
13. Caloi, L., Kotsakis, T., Palombo, M. R. & Petronio, C. *5th Int. Theriol. Congr. Abstr.* **1**, 141–142 (1989).
14. Stock, C. & Furlong, E. L. *Science* **68**, 140–141 (1928).
15. Roth, V. L. in *Body Size in Mammalian Paleobiology: Estimation and Biological Implications* (eds Damuth, J. & MacFadden, B. J.) 151–179 (Cambridge Univ. Press, Cambridge, 1990).
16. Vangengeim, E. A. *Trudy Geol. Inst. AN SSSR* (in Russian) **48**, 1–183 (1961).
17. Garutt, V. E. in *Anthropogene Period in the Arctic and Subarctic* (in Russian) 106–130 (Nedra, Moscow, 1965).
18. Lister, A. M. *Nature* **342**, 539–542 (1989).
19. Hopkins, D. M. *Quat. Res.* **3**, 520–540 (1973).
20. Sher, A. V. in *The 22nd Arctic Workshop 1992* 125–127 (University of Colorado, Boulder, 1992).
21. Petrovskii, V. V. & Yurtsev, B. A. in *The Arctic Ocean and its Coast in the Cenozoic Era* (in Russian) 509–515 (Gidrometeoizdat, Leningrad, 1970) (in English: Amerind Publishing, New Delhi, 1982).
22. Yurtsev, B. A. in *Paleoecology of Beringia* (eds Hopkins, D. M., Matthews, J. V. Jr, Schweger, C. E. & Young, S. B.) 157–177 (Academic, New York, 1982).
23. Sher, A. V. & Garutt, V. E. *Trudy Zool. Inst. AN SSSR* (in Russian) **131**, 93–103 (1977).
24. Haynes, G. *Mammoths, Mastodons, and Elephants: Biology, Behavior, and the Fossil Record* (Cambridge Univ. Press, Cambridge, 1991).
25. Arslanov, Kh. A. *Radiocarbon: Geochemistry and Geochronology* (in Russian) (Leningrad State Univ. Press, Leningrad, 1987).
26. Kiryushina, M. T. in *Anthropogene Period in the Arctic and Subarctic* (in Russian) 89–105 (Nedra, Moscow, 1965).
27. Garutt, V. E. *Trudy Zool. Inst. AN SSSR* (in Russian) **73**, 3–36 (1977).
28. Sher, A. V. & Garutt, V. E. *Dokl. AN SSSR* **285**, 221–225 (1985).
29. Garutt, V. E. & Foronova, I. V. *Research of the Teeth of Extinct Elephants: Methodical Recommendations* (in Russian) (Inst. Geol. i Geofiz. SO AN SSSR, Novosibirsk, 1976).
30. Lister, A. M. & Joysey, K. A. in *Structure, Function and Evolution of Teeth* (eds Smith, P. & Tchernov, E.) 185–213 (Freund, Jerusalem, 1992).
31. Sher, A. V. in *Beringia in the Cenozoic Era* (in Russian) 227–241 (USSR Academy of Sciences, Far-Eastern Scientific Center, Vladivostok, 1976) (in English: Amerind Publishing, New Delhi, 1984).

ACKNOWLEDGEMENTS. We thank T. V. Tertychnaya and S. B. Chernov for analytical work to date the samples; L. D. Sulerzhitsky for unpublished <sup>14</sup>C data and discussions; B. A. Yurtsev and V. Yu. Razzhin for help; V. Eisenmann for editing the English text; and A. M. Lister for useful comments and support. The field research has been supported by Polar Mercury (Pevke, Chukotka) and its Director, S. L. Girbasov, and by the Wrangel Island State Reserve.

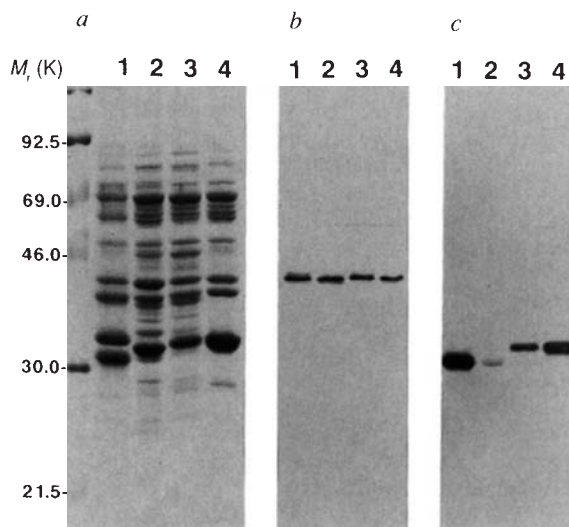


572 667

B. b B31 CAGCTAAATGTTGCAAACTCTTTTCTCTGGTGAGGGAGCT CAAACTGCTCAGGCTGCACCGGTTCAAGAGGGTGTTCACAGGAAGGAGCTCAACAGC  
ACAI -----TG----- --G-----T-----C-----G-A-----G-A-  
NBS16 -T-----A-----A----- --GG-----A-----T-----A-A-C-----T-----T-A-  
IUB18 -T-----A-----A----- --GG-----A-----T-----A-A-C-----A-----T-----T-A-  
-----TG-A-----A-T----- --GG-----CA-----AGA-GG -GC-CAA -----G-----G-TGCAAGC-ACT-  
B. c -----T-A-----T-CAA--GTA--CA-----AGA-GG TGC-CA-----G-A-G-CACAAGC-GCT-  
B. b -----T-----CA-----AGA-A-----G-CA----- --G-----TCAAGC-GCT-

by the dideoxy chain-termination method<sup>16</sup>. Primers used for PCR amplification and nucleotide sequencing were derived from the flagellin gene sequence<sup>9</sup>; the localization and sequence of primers are shown.

The ecology of *B. burgdorferi* is complex and involves a number of mammals, birds, reptiles and blood-sucking arthropods<sup>3</sup>. Its main vectors are ticks in the *I. ricinus* complex



341

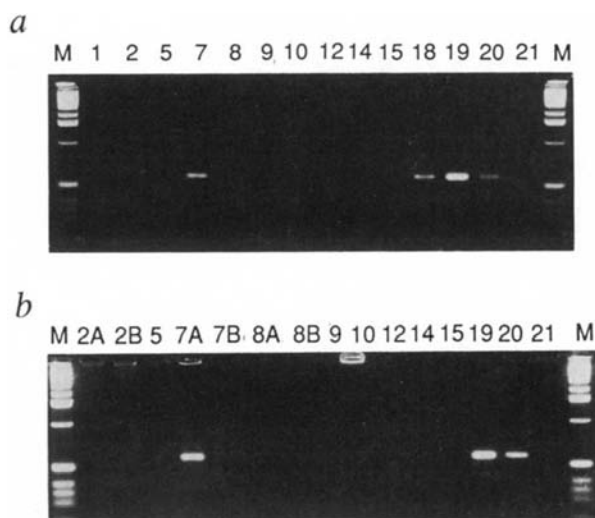


FIG. 4 PCR amplification of total DNA extracted from foot web biopsies and from *I. uriae* using the *B. burgdorferi* *fla* gene specific primers depicted in Fig. 2a, DNA extracted from biopsies obtained from the foot web of nine razorbill fledgelings (lanes 1–15) collected at Bonden island in July 1992; b, DNA extracted from *I. uriae* collected from the razorbill fledgelings investigated in a. Birds 2, 7 and 8 carried two ticks each (A and B) and bird 1 carried none. Lane 18 in a is the amplification product obtained using DNA isolated from the positive *B. burgdorferi* IUB18 culture. Lane 19, amplification of B31 DNA; lane 20, amplification of ACAI DNA; and lane 21, amplification of *Saccharomyces cerevisiae* DNA. For PCR amplification, the manufacturer's (Perkin Elmer, Cetus) standard protocol was followed. Running conditions were as follows: 35 cycles with a denaturation temperature of 94 °C for 2 min, annealing at 45 °C for 1 min, and elongation at 72 °C for 1 min. Oligomers used for flagellin gene amplification are shown in Fig. 2. M, 1-kilobase ladder (BRL).

(subgenus *Ixodes*). In this study we found that *I. uriae* (subgenus *Ceratixodes*) can also harbour *B. burgdorferi* s.l.. Using PCR, we showed that biopsies from razorbill nestlings contained *B. burgdorferi* DNA. Thus, it is very likely that the birds are infected with this spirochaete. But, we have not yet experimentally proved that razorbills are infective against ticks and competent reservoirs of *B. burgdorferi*. In contrast to most other ticks associated with *B. burgdorferi*, *I. uriae* at Bonden seems to have a narrow host range. The spirochaetes infecting *I. uriae* at Bonden are identical with *B. burgdorferi* from *I. ricinus* from the nearby island of Norrbyskär, indicating that there must be a close relationship between these populations of *B. burgdorferi*. We believe that birds on migration reach the island and may transmit *B. burgdorferi* between the different *Ixodes* species. The importance of birds as the major vertebrate reservoir for *B. burgdorferi* and of *I. uriae* as the primary vector at Bonden is evident. As auks spend most of the year far away from land and no mammals are found on the island, it is likely that at Bonden these birds are the main or only hosts for *I. uriae* and the only vertebrate reservoir for the borreliae found. It also seems clear that the spirochaetes in *I. uriae* at Bonden take part in a bird-tick cycle, less complex than those (involving mammals) previously described. Our results show that seabirds may be important in the maintenance and dissemination of *B. burgdorferi*. *I. uriae* has been recorded from more than 48 different sea bird species both from the northern and southern hemispheres<sup>15</sup>. We have also found spirochaete-like organisms in *I. uriae* from seabirds trapped on Iceland (unpublished results). The relatively broad host range of *I. uriae* and its wide geographic distribution suggests that hitherto unknown foci of *B. burgdorferi* may exist. □

- Clifford, C. M. in *Arctic and Tropical Arboviruses* (ed. Kurstak, E.) 83–100 (Academic, New York, 1979).
- Olsson, B. & Grenmyr, V. *Vår Fågelvärld* **3**, 11–14 (1992).
- Pokki, J. *Acta zool. Fenn.* **164**, 13–17 (1981).
- Southwood, T. R. E. in *Ecological Methods with Particular Reference to the Study of Insect Populations* 170–201 (Chapman and Hall, New York, 1978).
- Filippova, N. A. *Ixodid Ticks Subfamily Ixodinae Fauna SSSR* (Nov. Ser. no **114**, Paukoobraznye 4, Nauka, Moscow, 1977).
- Barbour, A. G., Hayes, S. F., Heiland, R. A., Schrupf, M. E. & Tessler, S. L. *Infect. Immun.* **52**, 549–554 (1986).
- Gassman, G. S., Kramer, M., Göbel, U. B. & Wallich, R. *Nucleic Acids Res.* **17**, 3590 (1989).
- Bergström, S., Bundoc, V. G. & Barbour, A. G. *Molec. Microbiol.* **3**, 479–486 (1989).
- Barbour, A. G. *Yale J. Biol. Med.* **57**, 521–525 (1984).
- Picken, R. *J. clin. Microbiol.* **30**, 99–114 (1992).
- Bergström, S. et al. *Scand. J. Infect. Dis.* **24**, 181–188 (1992).
- Barbour, A. G., Tessler, S. L. & Todd, W. J. *Infect. Immun.* **41**, 795–804 (1983).
- Kolonin, G. V. *World Distribution of Ixodid Ticks (Genus Ixodes)* (Nauka, Moscow, 1981).
- Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463–5467 (1979).

ACKNOWLEDGEMENTS. We thank G. Aronsson for technical assistance, A. G. Barbour for monoclonal antibodies H9724 and H5332, and B. Frankel, A. Tärnvik and N. Burman for critically reading the manuscript. This work was supported by grants to S.B. from the Swedish Medical Research Council, the Swedish Council for Forestry and Agricultural Research, Swedish Research Council for Engineering Sciences, the Medical Faculty of the University of Umeå, the J. C. Kempes Foundation, and Hierta Retzius Foundation; and by grants to T.G.T.J. from the Swedish Natural Science Research Council and Magnus Bergvalls Stiftelse.

## The neural correlates of the verbal component of working memory

E. Paulesu\*†, C. D. Frith\*‡§ & R. S. J. Frackowiak\*

\*MRC Cyclotron Unit, Hammersmith Hospital, DuCane Road, London W12 0HS, UK

†Istituto Scientifico S. Raffaele, via Olgettina 60, 20132 Milano, Italy

‡Psychological Department, University College London, Gower Street, London WC1E 6BT, UK

By repeating words 'in our head', verbal material (such as telephone numbers) can be kept in working memory<sup>1</sup> almost indefinitely. This 'articulatory loop' includes a subvocal rehearsal system<sup>2–6</sup> and a phonological store<sup>3,6–10</sup>. Little is known about neural correlates of this model of verbal short-term memory. We therefore measured regional cerebral blood flow, an index of neuronal activity, in volunteers performing a task engaging both components of the articulatory loop (short-term memory for letters)<sup>5–10</sup> and a task which engages only the subvocal rehearsal system (rhyming judgement for letters)<sup>4,11</sup>. Stimuli were presented visually and the subjects did not speak. We report here that comparisons of distribution of cerebral blood flow in these conditions localized the phonological store to the left supramarginal gyrus whereas the subvocal rehearsal system was associated with Broca's area. This is, to our knowledge, the first demonstration of the normal anatomy of the components of the 'articulatory loop'.

Our paradigm capitalizes on the observation that visually presented letters are transformed into a phonological code and access the articulatory loop, through the subvocal rehearsal system, even when there is no direct auditory stimulation<sup>5,6,10</sup>. In our experimental tasks, English letters were presented on a monitor, while in the control tasks, which were formally identical, we displayed Korean letters, which could not be transcoded phonologically. The 'Korean' versions included visual but not phonological components of the experimental tasks, and therefore provided a baseline for comparison (see Fig. 1 for details). In the first experimental task a string of letters was remembered, engaging all components of the articulatory loop, including the phonological store<sup>5–10</sup>. In the second experimental task, rhyming judgements were made. When letters are presented visually, rhyming decisions engage the subvocal rehearsal system<sup>4,5,12</sup>, but not the phonological store<sup>4,11</sup>.

Received 2 December 1992; accepted 20 January 1993.

- Anderson, J. F., Johnson, R. C., Magnarelli, L. A. & Hyde, F. W. *Infect. Immun.* **51**, 394–396 (1986).
- Anderson, J. F. *Rev. Infect. Dis.* **11** (suppl. 6), 1451–1458 (1989).

§ To whom correspondence should be addressed.



Six European male right-handed normal volunteers underwent 12 consecutive regional cerebral blood flow (rCBF) measurements (three for each experimental and control condition) with positron emission tomography (PET). We adopted a fully factorial design to compare the distributions of rCBF using statistical parametric mapping (SPM)<sup>13,14</sup>. Initially, we examined the main effect of phonological processing by combining the results from the two experimental tasks and comparing

them with the control tasks. This analysis resulted in a map of cerebral areas in which differences in rCBF occurred. Significant changes were identified by applying a statistical threshold of 0.05 corrected for multiple comparisons<sup>14</sup> (see Fig. 2 legend for methods). Table 1 reports stereotactic coordinates of significant rCBF increases due to phonological processing (statistical maps and cortical renderings are shown in Fig. 2). Significant activation foci were detected bilaterally in Brodmann's area (BA) 44,

FIG. 1 Schematic layout illustrating the cognitive components involved in each task (dark squares). In a prior experiment, we demonstrated that phonological coding did not occur with Korean letters. Six subjects did 60 trials of the short-term memory tasks for visually presented letters with or without concurrent articulatory suppression. Articulatory suppression disrupted performance (expressed as percentage of correct answers) in the English version of the task (mean difference: 13%; paired *t* test:  $t=8.9$ ; d.f.=5;  $P<0.0003$ ) but not in the Korean version (mean difference: 0.005%;  $t=0.28$ ; d.f.=5; n.s.). Experiment 1: Phonological short-term memory (experimental) task. Randomized sequences of six phonologically dissimilar consonants were presented on a computer screen at the rate of 1 per second. Subjects were instructed to rehearse the stimuli silently and to remember them. Two seconds after each sequence, a probe consonant appeared and subjects judged if it was present in the previous sequence. Subjects responded pointing with a joy-stick to yes/no symbols. Visual short-term memory (control) task. This task was identical to the previous one with the exception that Korean letters (not phonologically codeable) were used. Subjects were asked to remember the stimuli using a visual code. Experiment 2: Phonological-similarity (rhyming) (experimental) task. Subjects were asked to make rhyme judgments about consonants appearing at a rate of one per second. They moved a joy-stick towards a 'yes' symbol every time a letter appeared that rhymed with the letter 'B' which was always present on the screen. Rhyming letters occurred at a frequency of 1 in 6. Shape-similarity (control) task. Subjects were asked to judge whether a Korean letter looked similar to a Korean target letter always present on the screen.

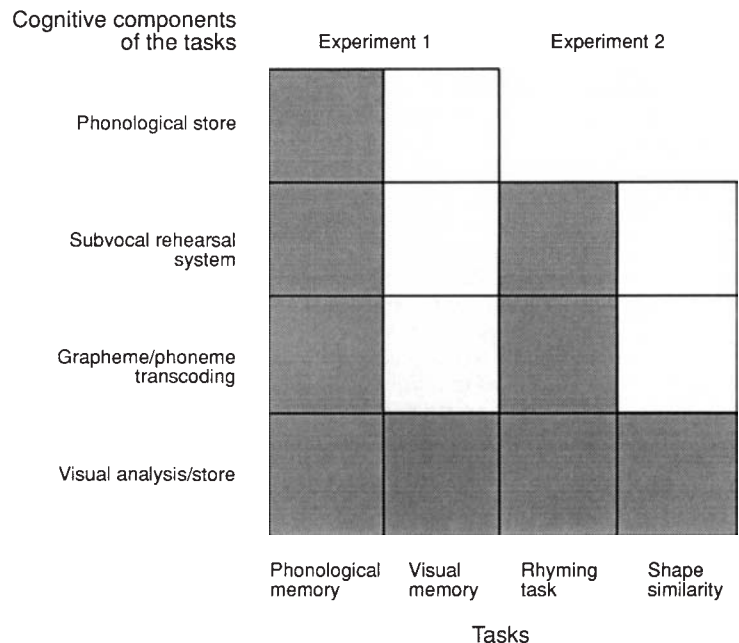
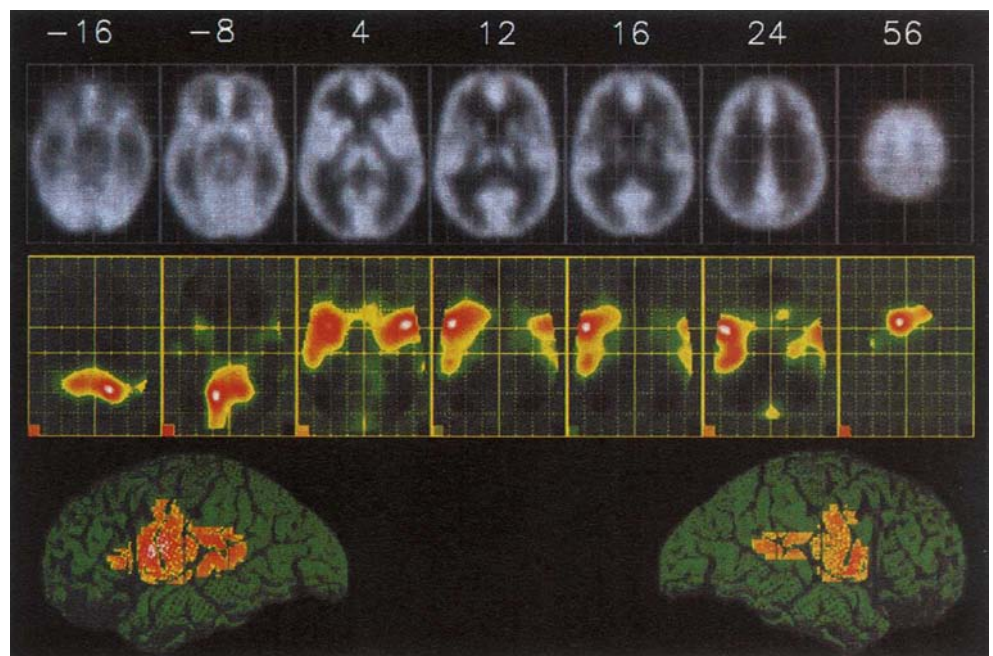


FIG. 2 Cerebral regions associated with phonological processing. The top row illustrates average blood flow maps from all subjects for all tasks in transaxial slices in the stereotactic anatomical space<sup>20</sup>. The location of rCBF increases associated with phonological processing is shown as *t*-maps in the second row. Numbers refer to the distance of the transaxial images from the intercommissural (AC-PC) level. The exact coordinates of the activated areas and their significance levels are given in Table 1. These areas of activation have been rendered onto the lateral surface of the cortex of the brain used by Talarach and Tournoux<sup>20</sup> to illustrate the distribution of the major sites of activation (left hemisphere on the left).

**METHODS.** rCBF was measured by recording the distribution of radioactivity following the intravenous injection of <sup>15</sup>O-labelled water (H<sub>2</sub><sup>15</sup>O) with the CTI 953B PET scanner (CTI Inc., Knoxville, TN, USA). Data were acquired by scanning in three-dimensional mode<sup>21</sup>. H<sub>2</sub><sup>15</sup>O was infused (10 ml min<sup>-1</sup>; 55 MBq ml<sup>-1</sup>) as a tracer of blood flow while scans were acquired. After attenuation correction (measured by a transmission scan), the data were reconstructed as 31 transaxial planes by three-dimensional filtered back projection with a Hanning filter of cutoff frequency 0.5 cycles per pixel. The resolution of the resulting images was 8.5×8.5×4.3 mm at full width at half-maximum<sup>22</sup>. The integrated counts accumulated were used as an index of rCBF<sup>23</sup>. The original brain images were transformed into a standard stereotactic anatomical space<sup>20,24</sup>. By



using appropriate statistical contrasts the main effect of phonological processing was estimated on a pixel by pixel basis using statistical parametric mapping (SPM). Global differences in CBF were first covaried out<sup>13</sup>, and comparisons of the means were made using the *t* statistic<sup>13,14</sup>. The resulting set of *t* values constituted a statistical parametric map [SPM(*t*)]<sup>13,14</sup>.

superior temporal gyri (BA 22/42), supramarginal gyri (BA 40) and insulae. We propose that this assembly of areas constitutes the functional anatomy of the 'articulatory loop'.

Cerebral areas thought to be devoted to motor aspects of speech planning and execution (for example supplementary motor area (SMA), cerebellum and possibly primary sensorimotor areas of mouth and larynx (BA 4, 3, 2, 1) and BA 6) were activated by our experimental tasks even though there was no overt speech<sup>15</sup>. A further activation focus was found in the left BA 18 (lingual gyrus). This structure is frequently damaged in pure alexia<sup>16</sup> and we therefore speculate that it is associated with visual processing of letters.

We then attempted to separate the functional anatomy of the subvocal rehearsal system from the phonological store. We compared the magnitude of activation due to the phonological short-term memory task (remembering letters) with that of the rhyming task because experimental evidence suggests that the phonological store has a fundamental role in phonological short-term memory<sup>7-10</sup>, but not in rhyming<sup>4,11</sup> (see Fig. 3 legend for details). This analysis demonstrated that activation detected in the left BA 40 was associated with the phonological short-term memory task but not with the rhyming task, and hence, following the psychological model summarized in Fig. 1, we believe that left BA 40 is the primary location of the phonological store.

The interpretation of our experiment is also corroborated by results of other studies of the functional anatomy of phonological processing. Passive perception of phonemes activated BA 22/42 bilaterally<sup>17</sup>, whereas a rhyming task<sup>17</sup> and a similar phonological judgement task<sup>18</sup>, all performed silently, also activated Broca's area (left BA 44) but not the left supramarginal gyrus (left BA 40). BA 22/42 is therefore probably involved in phonological processing independent from memory<sup>17</sup> and our data demonstrate that it can be activated even in the absence of external auditory phonological stimulation.

Our results indicate that Broca's area is crucial to the subvocal rehearsal system because both our experimental tasks involved this system<sup>3,5,12</sup> and both strongly activated BA 44 particularly in the left hemisphere (Fig. 3). The activation of SMA, cerebellum and possibly of sensory-motor areas suggests that these structures were automatically engaged as part of a more general neuronal network devoted to language planning and execution.

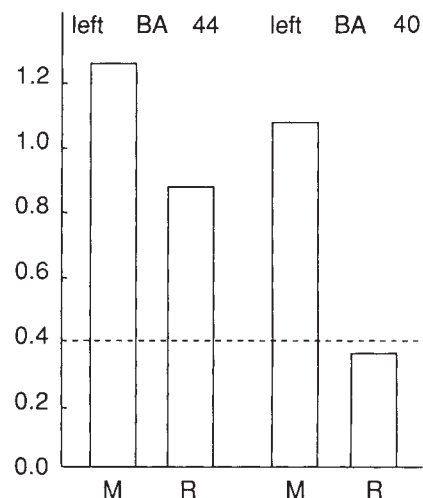


FIG. 3 Comparisons of experimental tasks. Mean increases (rCBF equivalents: ml per 100 g per min) in Broca's area (left BA 44) and in the left supramarginal gyrus (left BA 40) observed for the phonological short-term memory task (M) and the rhyming task (R) compared with their control tasks. Integrated count images were first converted into rCBF equivalents and normalized to a mean CBF of 50 ml per 100 g per min. The horizontal dashed line represents the 95% confidence limit for rCBF changes. The stereotactic locations are those reported in Table 1 for left BA 44 and BA 40. Broca's area was strongly activated in both the phonological short-term memory task and the rhyming task, while left BA 40 was activated by the phonological short-term memory task only. This was the only area where there was a significant interaction between the two experimental conditions (Z score, 4.2). We treated our two experiments as a fully factorial design with two treatments: (1) the nature of the stimuli and (2) the nature of the task. Each treatment had two levels: phonological and non-phonological stimuli for 1, the presence or the absence of loading in short-term memory for 2. The relative rCBF increase in the phonological short-term memory task compared with the rhyming task was assessed in terms of the interaction between treatments using linear contrasts.

Although it cannot be ruled out that some unconscious oropharyngeal muscle movement was occurring during the experimental tasks, 'inner speech' is probably not dependent on such movements. Indeed, anarthric patients, whose most common lesional correlate is a bilateral Rolandic operculum lesion including BAs 6, 4, 3, 2, 1, have a normally functioning articulatory loop despite being unable to produce overt speech<sup>5,19</sup>.

The novel attribution of left BA 40 to the phonological store on the basis of changes in cerebral blood flow supports previous anatomo-clinical correlation studies in brain-damaged patients. These studies, although not conclusive because of the variability in lesion site and size and the limited number of patients studied, suggested the left inferior parietal lobule, of which left BA 40 is part, as the location of the phonological store<sup>10</sup>. To our knowledge, this is the first account of the normal anatomy of the articulatory loop which also indicates the cerebral organization of the discrete processing and storage components. Our data support a multicomponent model for the articulatory loop, in which the subvocal rehearsal system and the phonological store are anatomically distinct. □

TABLE 1 Activation foci associated with phonological processing

| Region           | x (mm) | y (mm) | z (mm) | Z score    |
|------------------|--------|--------|--------|------------|
| SMA left         | -6     | 6      | 56     | 4.7        |
| SMA right        | 4      | 4      | 48     | 4.8        |
| BA 44 left*      | -46    | 2      | 16     | <b>9.0</b> |
| BA 44 right*     | 48     | 4      | 12     | 6.4        |
| BA 40 left       | -44    | -32    | 24     | <b>6.4</b> |
| BA 40 right      | 54     | -32    | 24     | 4.5        |
| BA 22/42 left    | -46    | -32    | 16     | <b>6.6</b> |
| BA 22/42 right   | 50     | -28    | 16     | 4.9        |
| BA 18 left       | -12    | -66    | -8     | <b>4.8</b> |
| Insula left      | -34    | 2      | 4      | 6.4        |
| Insula right     | 40     | 4      | 4      | 6.9        |
| Cerebellum left  | -18    | -54    | -16    | 4.1        |
| Cerebellum right | 14     | -60    | -16    | <b>5.5</b> |

Stereotactic coordinates refer to the maximal activation indicated by the highest Z score in a particular cerebral structure. Distances refer to the stereotactic space defined by Talairach and Tournoux<sup>20</sup>. Z scores in bold indicate structures in which activity was asymmetric. SMA, supplementary motor area; BA, Brodmann's area.

\* The site of maximally changed local blood flow was found to be in area 44. The extent as seen on the rendered cortical surface (Fig. 2) appears to include BA 6, 4, 3, 2, 1 bilaterally and BA 45 on the left. But in the absence of a separate discrete focus of change, no formal functional attribution to these cortical regions is valid.

Received 29 September 1992; accepted 14 January 1993.

1. Baddeley, A. D. *Science* **255**, 556-559 (1992).
2. Baddeley, A. D., Thomson, N. & Buchanan, M. *J. Verb. Learn. Verb. Behav.* **14**, 575-589 (1975).
3. Baddeley, A. D., Lewis, V. & Vallar, G. *Q. J. exp. Psychol.* **36**, 233-252 (1984).
4. Burani, C., Vallar, G. & Bottini, G. *Eur. J. cogn. Psychol.* **3**, 379-398 (1991).
5. Vallar, G. & Cappa, S. F. *Cogn. Neuropsychol.* **4**, 55-78 (1987).
6. Salamé, P. & Baddeley, A. D. *J. Verb. Learn. Verb. Behav.* **21**, 150-164 (1982).
7. Warrington, E. K. & Shallice, T. *Brain* **92**, 885-896 (1969).
8. Shallice, T. & Butterworth, B. *Neuropsychologia* **15**, 729-735 (1977).



9. Vallar, G. & Baddeley, A. D. *J. Verb. Learn. Verb. Behav.* **23**, 151–161 (1984).
10. Shallice, T. & Vallar, G. in *Neuropsychological Impairments of Short-term Memory* (eds Shallice, T. & Vallar, G.) 11–53 (Cambridge Univ. Press, New York, 1990).
11. Vallar, G. & Baddeley, A. D. *Cogn. Neuropsychol.* **1**, 121–141 (1984).
12. Besner, D. Q. *J. exp. Psychol.* **39A**, 467–478 (1987).
13. Friston, K. J. *et al. J. Cereb. Blood Flow Metab.* **10**, 458–466 (1990).
14. Friston, K. J., Frith, C. D., Liddle, P. F. & Frackowiak, R. S. J. *J. Cereb. Blood Flow Metab.* **11**, 690–699 (1991).
15. Wise, R. *et al. Brain* **114**, 1803–1817 (1991).
16. Damasio, A. R. & Damasio, H. *Neurology*, **33**, 1573–1583 (1983).
17. Zatorre, R. J., Evans, A. C., Meyer, E. & Gjedde, A. *Science* **256**, 846–849 (1992).
18. Démonet, J. F. *et al. Brain* **115**, 1753–1768 (1992).
19. Baddeley, A. & Wilson, B. *J. Mem. Lang.* **24**, 490–502 (1985).
20. Talairach, J. & Tournoux, P. *A Co-planar Stereotaxic Atlas of the Human Brain* (Thieme, Stuttgart, 1988).
21. Townsend, D. W. *et al. IEEE Trans. med. Imaging* **10**, 505–512 (1991).
22. Spinks, T. J. *et al. Phys. med. Biol.* **37**, 1637–1655 (1992).
23. Fox, P. T. & Mintun, M. A. *J. nucl. Med.* **30**, 141–149 (1989).
24. Friston, K. J. *et al. J. Cereb. Blood Flow Metab.* **9**, 690–695 (1989).

ACKNOWLEDGEMENTS. We thank G. Lewington and L. Schnorr for their help in scanning the volunteers, our colleagues in the MRC Cyclotron Unit Chemistry and PET Physics Sections for making this research possible and C. J. Bench, G. Bottini, J. F. Démonet, P. M. Grasby and R. E. Passingham for discussion. The studies were approved by the Hammersmith Hospital Medical Ethics Committee and permission to administer radioactivity was obtained from the ARSAC, UK.

## Different binding epitopes on the NK<sub>1</sub> receptor for substance P and a non-peptide antagonist

Ulrik Gether\*, Teit E. Johansen\*†, R. Michael Snider‡, John A. Lowe III‡, Shigetada Nakanishi§ & Thue W. Schwartz\*

\* Laboratory of Molecular Endocrinology, University Department of Clinical Biochemistry, Rigshospitalet 6321, Blegdamsvej 9, DK-Copenhagen 2100, Denmark

† Department of Exploratory Medicinal Chemistry, Central Research Division, Pfizer Inc., Groton, Connecticut 06340, USA

§ Institute for Immunology, Faculty of Medicine, Kyoto University, Kyoto 606, Japan

**NON-PEPTIDE** ligands for peptide receptors have been discovered in several systems through file screening programs<sup>1–6</sup>, but the mechanism of action for these candidate drugs is obscure as they do not chemically resemble the native peptides. The compound CP 96345 is a high-affinity, non-peptide antagonist of the substance P (NK<sub>1</sub>) receptor<sup>4,5,7</sup>, which is important in pain perception and neurogenic inflammation<sup>8–11</sup>. Here we identify epitopes on the NK<sub>1</sub> receptor responsible for the specific binding of CP 96345 by systematic exchange of corresponding segments between the NK<sub>1</sub> receptor and the homologous NK<sub>3</sub> (neurokinin B) receptor, which does not bind the non-peptide ligand. Non-conserved residues, in two epitopes around the top of transmembrane segment V and in one epitope at the top of transmembrane segment VI, are essential for the specific action of CP 96345 on the NK<sub>1</sub> receptor, but are surprisingly not important for the binding of the natural peptide ligand, substance P. Susceptibility to the non-peptide antagonists can be conveyed to the previously unresponsive NK<sub>3</sub> receptor by mutational transfer of this discontinuous epitope from the NK<sub>1</sub> receptor.

CP 96345 potentially inhibited binding of radiolabelled substance P to rat cloned NK<sub>1</sub> receptor ( $K_i = 14$  nM) but does not react at all with the homologous cloned rat NK<sub>3</sub> (neurokinin B) receptor ( $K_i > 0.01$  mM) (Table 1). To identify regions of the NK<sub>1</sub> receptor responsible for the high-affinity binding of CP 96345, we therefore systematically exchanged segments of the NK<sub>1</sub> receptor<sup>12</sup> with corresponding segments from the NK<sub>3</sub> receptor<sup>13</sup> (Fig. 1 and Table 1). The general structural integrity of the chimaeric receptors could be demonstrated by the amphibian, non-selective tachykinin peptide, elodeisin, which

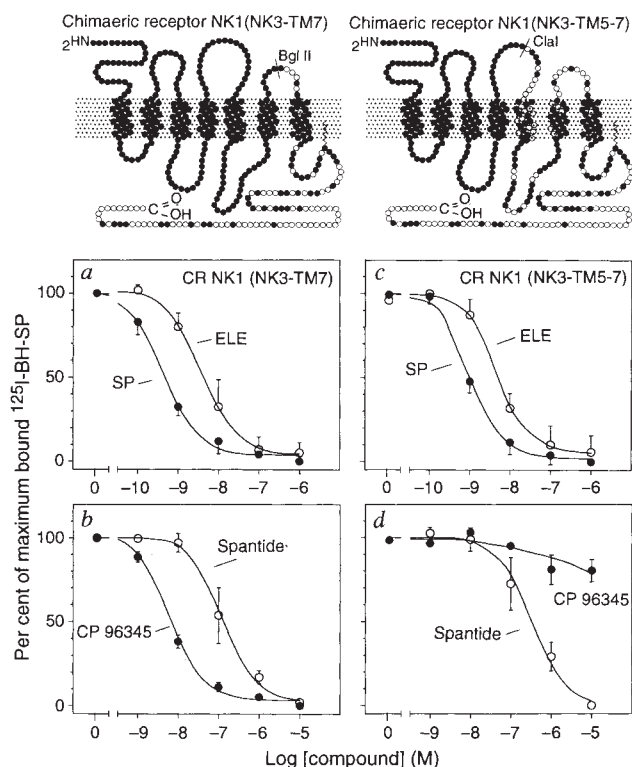


FIG. 1 Identification of domains involved in CP 96345 action using chimaeric NK<sub>1</sub>/NK<sub>3</sub> receptor constructs. Top panel, structure of CR NK1(NK3-TM7) and CR NK1(NK3-TM5-7). Black circles indicate amino acids in the NK<sub>1</sub> receptor or conserved amino-acid residues between the NK<sub>1</sub> and NK<sub>3</sub> receptor. Open circles indicate amino acids specific for the NK<sub>3</sub> receptor. The chimaeric splicing junctions are indicated. Middle panel, competition binding of substance P (●) and elodeisin (○) with <sup>125</sup>I-BH-SP to CR NK1(NK3-TM7) (a) and CR NK1(NK3-TM5-7) (c). Bottom panel, competition binding of CP96345 and spantide with <sup>125</sup>I-BH-SP to CR NK1(NK3-TM7) (b) and CR NK1(NK3-TM5-7) (d). Binding was assayed on intact cells as described in the legend to Table 1. Data are expressed as mean ± s.e.m.,  $n = 3-5$ . Note different scales on the x-axis.

bound with similar affinity to both wild-type and chimaeric receptors (Table 1).

The ability of CP 96345 to inhibit binding of radiolabelled substance P did not change when the transmembrane segment VII (TM VII) and the C-terminal tail of the NK<sub>1</sub> receptor were exchanged with the corresponding segments of the NK<sub>3</sub> receptor (chimaeric receptor (CR) NK1(NK3-TM7) (Fig. 1 and Table 1). But when TM V and VI and a part of the second extracellular loop were also exchanged (CR NK1(NK3-TM5-7), the ability of CP 96345 to inhibit radiolabelled substance P binding was eliminated ( $K_i > 0.01$  mM) (Fig. 1 and Table 1). This indicates that some amino-acid residues that differ between the NK<sub>1</sub> and NK<sub>3</sub> receptor in or around TM V and TM VI (open symbols in Fig. 1, top panel) are essential for the binding of CP 96345 to the NK<sub>1</sub> receptor. These residues appear not to be important for the selective binding of the natural ligand, substance P, which was unaffected by the exchange of receptor segments (Fig. 1 and Table 1). If residues around TM V and TM VI are involved in binding of substance P, our data indicate that such residues must be identical or similar in the NK<sub>1</sub> and NK<sub>3</sub> receptor.

The binding affinity of three peptide antagonists of substance P<sup>14–16</sup> dramatically dropped when both TM VII and VI were exchanged (Table 1). Surprisingly, and in contrast to CP 96345, both spantide ([D-Arg<sup>1</sup>, D-Trp<sup>7,9</sup>, Leu<sup>11</sup>]substance P) and [Arg<sup>6</sup>, D-Trp<sup>7,9</sup>, MePhe<sup>8</sup>]substance P(6–11) retained almost all their binding affinity for the chimaeric receptor in which both TM V and TM VI originated from the NK<sub>3</sub> receptor (Fig. 1 and

† Present address: Johns Hopkins University, Department of Medicine, Bayview Circle, Baltimore, Maryland 21224, USA.

Table 1). This may reflect that some degree of molecular incompatibility between TM V from the NK<sub>1</sub> receptor and TM VI from NK<sub>3</sub> receptor is perturbing the antagonist binding but not the agonist binding. It was also found in the original study on systematic chimaeric constructs between the adrenergic  $\alpha_2$  and  $\beta_2$  receptors that the binding of the antagonist radioligand was lost when TM V originated from  $\beta_2$  and TM VI from  $\alpha_2$  (ref. 17).

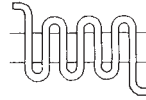
To characterize further the CP 96345-reactive epitope, we constructed four chimaeric NK<sub>1</sub> receptors substituted with small segments from the NK<sub>3</sub> receptor corresponding to the upper part of TM V ('B'), and TM VI ('C') and the adjacent parts of the extracellular loops ('A' and 'D') (Fig. 2). Substance P bound to all these mutated receptors with an affinity similar to that of the wild-type NK<sub>1</sub> receptor (legend to Fig. 2). The exchange of segment C did not affect the action of CP 96345, indicating that the three residues that differ between the NK<sub>1</sub> and NK<sub>3</sub> receptor in the upper part of TM VI do not appear to be involved in CP 96345 binding (Fig. 2). But the exchange of all the three other segments (A, B and D) decreased the susceptibility of the NK<sub>1</sub> receptor to CP 96345 in various degrees, with the residues in the A domain appearing to be most important (Fig. 2). These experiments indicate that the action of CP 96345 is dependent upon non-conserved residues in three subepitopes located at the junction between extracellular loop-2 and the top of TM V and at the top of TM VI, respectively.

To elucidate whether transfer of these three subepitopes (A, B and D) (Fig. 2) from the NK<sub>1</sub> receptor could confer CP 96345 binding ability on the previously unresponsive NK<sub>3</sub> receptor, we constructed the reciprocal chimaeric receptors (Fig. 3). Transfer of segment B from the NK<sub>1</sub> receptor resulted in a modest increase in affinity for CP 96345 compared with the wild-type NK<sub>3</sub> receptor, whereas transfer of segment A or D each conferred CP 96345 binding ability on the NK<sub>3</sub> receptor, albeit with low affinity (Fig. 3). Importantly, the combined transfer of segments A and D from the NK<sub>1</sub> receptor was able

to provide the NK<sub>3</sub> receptor with an affinity for CP 96345 ( $K_1 = 82 \pm 1$  nM; mean  $\pm$  s.e.m.,  $n = 3$ ) that was only 6-fold lower than the affinity for the wild-type NK<sub>1</sub> receptor ( $K_1 = 14$  nM  $\pm$  6 nM; mean  $\pm$  s.e.m.,  $n = 3$ ) (Fig. 3). We could convey full NK<sub>1</sub>-like affinity for CP 96345 to the NK<sub>3</sub> receptor when we transferred the whole segment around TM V and VI from the NK<sub>1</sub> receptor ( $K_1 = 15.2 \pm 1.6$  nM; mean  $\pm$  s.e.m.,  $n = 3$ ) (Fig. 3). These data demonstrate that the region around TM V and VI determine the receptor selectivity of CP 96345 and they suggest that the region constitutes an important part of the binding domain for the non-peptide ligand. It is still possible, however, that CP 96345 could have some points of interaction with residues that are conserved between the NK<sub>1</sub> and NK<sub>3</sub> receptor located outside this region. The binding affinity for substance P was slightly increased for CR NK3(NK1-TM5-6) compared with the wild-type NK<sub>3</sub> receptor ( $K_1 = 63 \pm 3$  nM versus  $300 \pm 80$  nM), indicating that substance P may have weak interactions in this region.

Previously, chimaeric and other mutational studies aimed at identifying antagonist binding sites on G-protein-coupled receptors have been done exclusively on receptors for small classical transmitters such as adrenergic and muscarinic receptors<sup>18</sup>. These studies have shown that residues located within TM VI and VII are important for the selective recognition of adrenergic antagonists and of some muscarinic antagonists<sup>17-20</sup>. But the binding of such antagonists also involves interactions with residues located in other transmembrane helices several Ångströms down into the membrane<sup>21-23</sup> and these points of contact are often shared with the corresponding endogenous agonists to which the antagonists are chemically related<sup>22</sup>. Our data indicate that the mechanism of action of the non-peptide antagonist CP 96345 could differ from that of the antagonists for the small classical transmitters. In contrast to these, CP 96345 possibly acts through residues in the receptor located close to the outer surface of the membrane. Furthermore, it appears from

TABLE 1 Binding\* of peptide agonists, peptide antagonists and the non-peptide antagonist CP 96345 to the NK<sub>1</sub>, NK<sub>3</sub> and chimaeric NK<sub>1</sub>/NK<sub>3</sub> receptors

|                                                                             | Wild type<br>NK-1                                                                   | CR NK1<br>(NK3-TM7)                                                                 | CR NK1<br>(NK3-TM6-7)                                                                | CR NK1<br>(NK3-TM5-7)                                                                 | Wild type<br>NK-3                                                                     |
|-----------------------------------------------------------------------------|-------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------|
|                                                                             |  |  |  |  |  |
| Substance P                                                                 | 0.14 $\pm$ 0.03                                                                     | 0.49 $\pm$ 0.04                                                                     | 1.1 $\pm$ 0.2                                                                        | 0.9 $\pm$ 0.1                                                                         | 300 $\pm$ 80                                                                          |
| Eledoisin                                                                   | 16 $\pm$ 5                                                                          | 4.6 $\pm$ 1.3                                                                       | 12 $\pm$ 3.5                                                                         | 4.5 $\pm$ 0.6                                                                         | 4.7 $\pm$ 1.1                                                                         |
| Spantide                                                                    | 100 $\pm$ 55                                                                        | 130 $\pm$ 60                                                                        | ~10,000                                                                              | 350 $\pm$ 100                                                                         | 8,000 $\pm$ 2,500                                                                     |
| [D-Pro <sup>2</sup> , D-Trp <sup>7,9</sup> , Ile <sup>11</sup> ]substance P | 170 $\pm$ 30                                                                        | 68 $\pm$ 6                                                                          | >10,000                                                                              | 8,500 $\pm$ 2,500                                                                     | >10,000                                                                               |
| [Arg <sup>6</sup> , D-Trp <sup>7,9</sup> , MePhe <sup>8</sup> ]substance P  | 700 $\pm$ 100                                                                       | 920 $\pm$ 160                                                                       | >10,000                                                                              | 830 $\pm$ 120                                                                         | >10,000                                                                               |
| CP 96345                                                                    | 14 $\pm$ 6                                                                          | 5.4 $\pm$ 0.9                                                                       | 330 $\pm$ 60                                                                         | >10,000                                                                               | >10,000                                                                               |

**Construction and expression of chimaeric receptors:** The chimaeric receptors between the rat NK<sub>1</sub> and NK<sub>3</sub> receptor were constructed using both pre-existing and introduced restriction sites located at equivalent positions in the receptor cDNAs<sup>12,13</sup>. Restriction sites were introduced by site-directed mutagenesis in the M13 system<sup>24</sup>. A *Bgl*II and a *Cla*I site were introduced in the NK<sub>3</sub> receptor at Gln 314 (base pair 937) and at Val 222 (base pair 663), respectively, and a *Bcl*I and a *Cla*I site were introduced in the NK<sub>1</sub> receptor at Met 250 (base pair 758) and at Ile 182 (base pair 544), respectively. The difference in amino-acid numbering is caused by the N-terminal extension of the NK<sub>3</sub> receptor<sup>12,13</sup>. All restriction sites were silently introduced except the *Cla*I site in the NK<sub>1</sub> receptor, which contained the conservative mutation NK<sub>1</sub> Asp 183  $\rightarrow$  Glu 183. All mutations were verified by DNA sequencing<sup>25</sup>. The appropriate restriction endonuclease fragments encoding the chimaeric constructs were cloned into the pTEJ-8 expression vector<sup>26</sup>. The NK<sub>1</sub> and chimaeric receptors were stably transfected into CHO cells using calcium phosphate precipitation<sup>26,27</sup>. The NK<sub>3</sub> receptor was transiently expressed in COS-7 cells as it is expressed only in a low-affinity form in CHO cells but in its native, high-affinity form in COS cells<sup>27</sup>. Control experiments with the NK<sub>1</sub> receptor and CR NK1(NK3-TM5-7) transiently expressed in COS-7 gave binding affinities similar to those resulting from stable expression in CHO cells. **Binding assay:** Binding was assayed on intact cells as described<sup>27</sup>. Transfected cells were incubated for 3 h at 4 °C with 50 pM radioligand and the indicated concentrations of unlabelled peptide or non-peptide compound (triplicate determinations) in 0.5 ml of a 50 mM Tris-HCl buffer, pH 7.4, containing 150 mM NaCl and 5 mM MnCl<sub>2</sub> and supplemented with 0.1% bovine serum albumin (Sigma), 100  $\mu$ g ml<sup>-1</sup> bacitracin (Sigma), 5  $\mu$ g ml<sup>-1</sup> leupeptin (Sigma) and 10  $\mu$ g ml<sup>-1</sup> chymostatin (Sigma). The moniodinated HPLC-purified radioligands, [<sup>125</sup>I]-Bolton Hunter labelled substance P ([<sup>125</sup>I]-BH-SP), used for the wild-type NK<sub>1</sub> and chimaeric receptors, and eledoisin ([<sup>125</sup>I]-BH-ELE), used for the wild-type NK<sub>3</sub> receptor, were prepared as described<sup>27</sup>. All binding data were analysed and IC<sub>50</sub> values determined by computerized nonlinear regression analysis.  $K_D$  values for binding of radiolabelled substance P and eledoisin were estimated from competition binding data with 10–12 different concentrations of the corresponding unlabelled peptide using the equation,  $K_D = IC_{50} - L$ .

\*  $K_1$  values are given as mean  $\pm$  s.e.m.,  $n = 3$ –5, in nM.  $K_1$  values were calculated from  $K_1 = IC_{50} / (1 + L/K_D)$ , where  $L$  is the concentration of free radioligand.



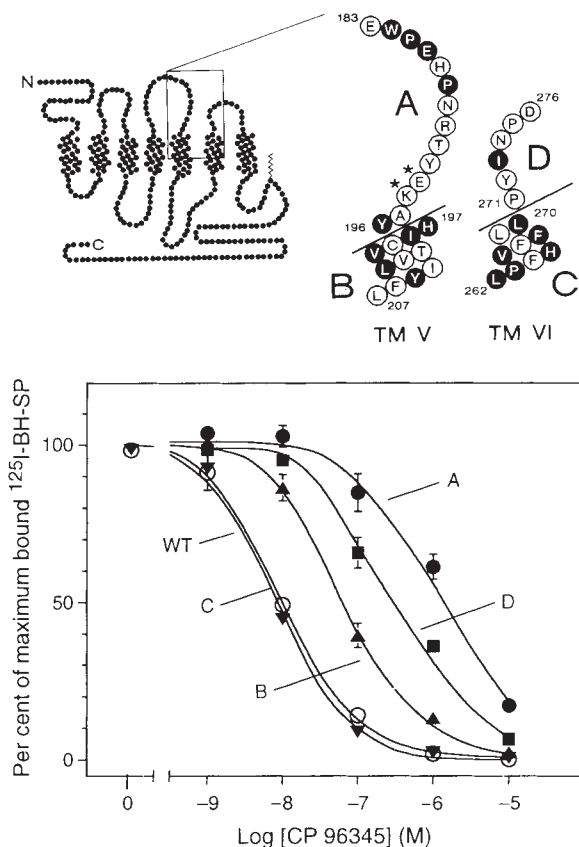


FIG. 2 Identification of epitopes around the top of TM V and VI of the  $NK_1$  receptor involved in CP 96345 action. Top panel, structure of the  $NK_1$  receptor in the region around TM V and TM VI. Black circles indicate amino acids conserved between the  $NK_1$  receptor and the  $NK_3$  receptor. Open circles indicate amino acids specific for the  $NK_1$  receptor. Four segments of the  $NK_1$  receptor were exchanged with corresponding segments from the  $NK_3$  receptor: the outer part of TM V, NK1(NK3/197-207) (B), and TM VI, NK1(NK3/262-270) (C), and the two adjacent parts of the extracellular loops, NK1(NK3/183-196) (A) and NK1(NK3/271-276) (D). The start of segment A corresponds to the junction between the two receptor fragments in CR NK1(NK3-TM5-7) and the end of segment D to the junction in CR NK1(NK3-TM7). The constructs, made by site-directed mutagenesis, were transiently expressed in COS-7 cells and binding was assayed as described in the legend to Table 1. Amino-acid residues marked with an asterisk are not present in the  $NK_3$  receptor and were detected as part of mutation A. Lower panel, competition binding of CP 96345 with  $^{125}$ I-BH-SP to the wild-type (wt)  $NK_1$  receptor (○) and to mutated  $NK_1$  receptors, A (●), B (▲), C (▼) and D (■). Data are expressed as mean  $\pm$  s.e.m.,  $n=3$ .  $K_D$  values for binding of  $^{125}$ I-BH-SP to the wild-type  $NK_1$  and mutated receptors transiently expressed in COS-7 cells were (mean  $\pm$  s.e.m.,  $n=3$ ): WT,  $K_D=0.27 \pm 0.04$  nM; A,  $K_D=0.5 \pm 0.2$  nM; B,  $K_D=0.15 \pm 0.03$  nM; C,  $K_D=0.25 \pm 0.06$  nM; D,  $K_D=0.18 \pm 0.05$  nM.

our results that CP 96345 acts, at least partly, through an epitope that is different from the substance P binding site. Both the non-peptide and peptide antagonists are dependent on epitopes around TM V and TM VI (Table 1). This suggests that the antagonistic mechanism could derive from interference with the correct packing of the adjacent helices. Interestingly, the intracellular loop that connects these two helices is crucial for G-protein coupling, signal transduction and the allosteric induction of the high-affinity state for this class of receptors<sup>17</sup>. This may be the basis of a general mechanism of action for non-peptide antagonists for peptide receptors.

The localization in the  $NK_1$  structure of a relatively small, discontinuous epitope, essential for the action of a non-peptide ligand, may make it possible to identify the exact points of chemical interaction by combining a series of systematic single substitutions in the binding epitope with a library of chemically

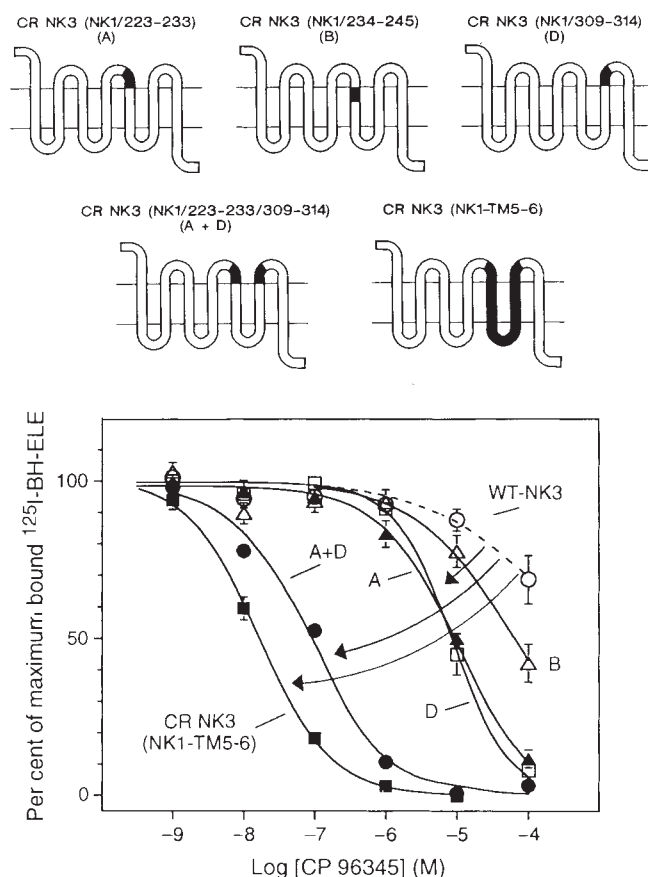


FIG. 3 Genetic transfer of epitopes around the top of TM V and VI of the  $NK_1$  receptor to the CP 96345 unresponsive  $NK_3$  receptor. Top panel, simplified structure of the  $NK_3$  receptor substituted with  $NK_1$  receptor segments as indicated by solid regions. The chimaeric receptors were constructed, transiently expressed in COS-7 cells and binding assayed as described in legend to Table 1. Bottom panel, competition binding of CP 96345 with  $^{125}$ I-BH-ELE to the wild-type  $NK_3$  receptor (○) and chimaeric receptors NK3(NK1/223-233) (A) (▲), NK3(NK1/234-245) (B) (△), NK3(NK1/309-314) (D) (□), NK3(NK1/223-233,309-314) (A+D) (●) and NK3(NK1-TM5-6) (■). Data are expressed as mean  $\pm$  s.e.m.,  $n=3-4$ .  $K_D$  values for binding of  $^{125}$ I-BH-ELE to the mutated receptors transiently expressed in COS-7 cells were (mean  $\pm$  s.e.m.,  $n=3-4$ ): NK3(NK1/223-233) (A),  $K_D=0.9 \pm 0.2$  nM; NK3(NK1/234-245) (B),  $K_D=2.4 \pm 0.8$  nM; NK3(NK1/309-314) (D),  $K_D=1.7 \pm 0.8$  nM; NK3(NK1/223-233,309-314) (A+D),  $K_D=0.78 \pm 0.01$  nM; NK3(NK1-TM5-6),  $K_D=4.7 \pm 1.6$  nM.

modified non-peptide analogues. This would facilitate future development of highly selective drugs acting on neuropeptide receptors. There are more than forty neuropeptides apart from the tachykinins that serve as chemical messengers throughout the human brain. The development of non-peptide compounds that can control these systems will have considerable therapeutic potential.

**Note added in proof:** In a recent study, published after submission of this manuscript and aimed at characterizing the binding site for substance P, mutations in the extracellular loop-2 of the  $NK_1$  receptor impaired the binding of an analogue of CP 96345 without affecting substance P binding, in agreement with our observations<sup>28</sup>. □

Received 21 October; accepted 30 December 1992.

- Freidinger, R. M. et al. in *Neuropeptides and their Receptors* (eds Schwartz, T. W., Hilsted, L. M. & Rehfeld, J. F.) 389-400 (Alfred Benzon Symposium 29, Munksgaard, Copenhagen, 1990).
- Chang, R. S. L. & Lotti, V. J. *Molec. Pharmac.* **29**, 347-351 (1989).
- Dudley, D. T. et al. *Molec. Pharmac.* **38**, 370-377 (1990).
- Snider, R. M. et al. *Science* **251**, 435-437 (1991).
- McLean, S. et al. *Science* **251**, 437-439 (1991).
- Yamamura, Y. et al. *Science* **252**, 572-574 (1991).

7. Snider, R. M., Longo, K. P., Drozda, S. E., Lowe III, J. A. & Leeman, S. E. *Proc. natn. Acad. Sci. U.S.A.* **88**, 10042-10044 (1991).
8. Pernow, B. *Pharmac. Rev.* **35**, 85-141 (1983).
9. Buck, S. H. & Burcher, E. *Trend. pharmac. Sci.* **7**, 65-68 (1986).
10. McGillis, P., Mitsuyoshi, M. & Payan, D. G. *Ann. N.Y. Acad. Sci.* **594**, 85-94 (1990).
11. Iversen, L. L., McKnight, A. T., Foster, A. C., Young, S. C. & Williams, B. J. in *Neuropeptides and their Receptors* (eds Schwartz, T. W., Histed, L. M. & Rehfeld, J. F.) 363-372 (Alfred Benzon Symposium 29, Munksgaard, Copenhagen, 1990).
12. Yokota, Y. *et al. J. biol. Chem.* **264**, 17649-17652 (1989).
13. Shigemoto, R., Yokota, Y., Tsuchida, K. & Nakanishi, S. *J. biol. Chem.* **265**, 623-628 (1990).
14. Engberg, G., Svensson, T. H., Rosell, S. & Folkers, K. *Nature* **293**, 222-223 (1981).
15. Mizrahi, J., Escher, E., D'Orleans-Juste, P. & Regoli, D. *Eur. J. Pharmac.* **99**, 193-202 (1984).
16. Laufer, R. *et al. Proc. natn. Acad. Sci. U.S.A.* **82**, 7444-7448 (1985).
17. Koblika, B. K. *et al. Science* **240**, 1310-1316 (1988).
18. Dohlman, H. G., Thorner, J., Caron, M. G. & Lefkowitz, R. J. *A. Rev. Biochem.* **60**, 653-688 (1991).
19. Frielle, T., Daniel, K. W., Caron, M. G. & Lefkowitz, R. J. *Proc. natn. Acad. Sci. U.S.A.* **85**, 9494-9498 (1988).
20. Wess, J., Gdula, D. & Brann, M. R. *Molec. Pharmac.* **41**, 369-374 (1992).
21. Dixon, R. A. F. *et al. Nature* **326**, 73-77 (1987).
22. Strader, C. D. *et al. J. biol. Chem.* **263**, 10267-10271 (1988).
23. Wess, J., Gdula, D. & Brann, M. R. *EMBO J.* **10**, 3729-3734 (1991).
24. Johansen, T. E., Vogel, C. & Schwartz, T. W. *Biochem. biophys. Res. Commun.* **172**, 1384-1391 (1990).
25. Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463-5467 (1977).
26. Johansen, T. E., Scheller, M. S., Tolstoy, S. & Schwartz, T. W. *FEBS Lett.* **267**, 289-294 (1990).
27. Gether, U., Marray, T., Schwartz, T. W. & Johansen, T. E. *FEBS Lett.* **296**, 241-244 (1992).
28. Fong, T. M., Huang, R. C. & Strader, C. D. *J. biol. Chem.* **267**, 25664-25667 (1992).

ACKNOWLEDGEMENTS. We thank M. R. Brann and R. Håkanson for comments on the manuscript, and T. Jakobsen, B. Friis, S. Hummelgaard and R. Hansen for technical assistance. The work was supported by grants from the Danish Medical Research Council, the NOVO Foundation, the Carlsberg Foundation and the Danish Biotechnology Center for Signal Peptides.

## A single amino acid of the cholecystokinin-B/gastrin receptor determines specificity for non-peptide antagonists

Martin Beinborn\*†, Young-Mee Lee\*,  
Edward W. McBride\*, Suzanne M. Quinn\*  
& Alan S. Kopin\*‡

\*Division of Gastroenterology and GRASP Digestive Disease Center, New England Medical Center, Tufts University School of Medicine, Boston, Massachusetts 02111, USA

†Department of General Pharmacology, Medical School, W-3000 Hannover 61, Germany

THE brain cholecystokinin-B/gastrin receptor (CCK-B/gastrin) has been implicated in mediating anxiety, panic attacks, satiety, and the perception of pain<sup>1-5</sup>. The canine and human CCK-B/gastrin receptors share 90% amino-acid identity<sup>6,7</sup> and have similar agonist affinities. These receptors can be selectively blocked by the non-peptide benzodiazepine-based antagonists L365260 (ref. 8) and L364718 (ref. 9); however, the binding of these antagonists to the human and canine receptors differs by up to 20-fold, resulting in a reversal of affinity rank order. Here we report the identification of a single amino acid in the sixth transmembrane domain of the CCK-B/gastrin receptor that corresponds to valine 319 in the human homologue and which is critical in determining the binding affinity for these non-peptide antagonists. We show that it is the variability in the aliphatic side chain of the amino acid in position 319 that confers antagonist specificity. Substitution of valine 319 with a leucine residue decreases the affinity for L365260 20-fold while concomitantly increasing the affinity for L364718. An isoleucine in the same position of the human receptor selectively increases affinity for L364718. Interspecies differences in the aliphatic amino acid occupying this single position selectively affect antagonist affinities without altering the agonist binding profile<sup>6,7,10</sup>. We therefore conclude that the residues underlying non-peptide antagonist affinity must differ from those that confer agonist specificity. To our knowledge, these findings are the first example in which a critical antagonist binding determinant for a

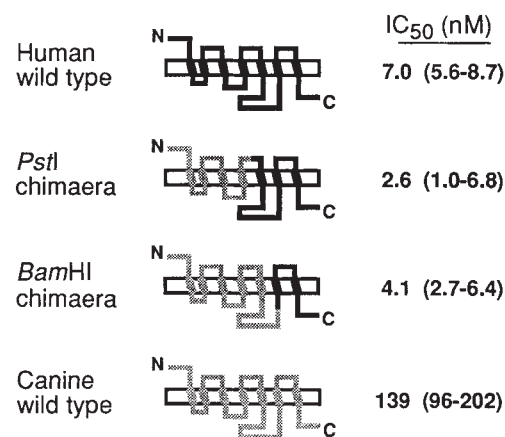


FIG. 1 Determinants for high-affinity binding of L365260 are found at the carboxy end of the human CCK-B/gastrin receptor. IC<sub>50</sub> values for L365260 were calculated from [<sup>125</sup>I] CCK-8 competition binding studies of recombinant wild-type and chimaeric CCK-B/gastrin receptors. cDNAs encoding the wild-type human and canine CCK-B/gastrin receptors were subcloned into the expression vector pcDNA 1 (Invitrogen). The *PstI* chimaera was constructed using the naturally occurring *PstI* sites which are found in the same relative position in both receptors (in human, nucleotides 607-612, and canine, nucleotides 619-624). Construction of the *BamHI* chimaera required the introduction of a *BamHI* restriction site (nucleotides 970-975) into the canine cDNA (see mutant canine G316-S in Fig. 2) in the position corresponding to a natural *BamHI* site in the human cDNA (nucleotides 943-948). Chimaeric receptor cDNA sequences were confirmed using a combination of dideoxy nucleotide sequencing and restriction endonuclease analysis. Wild-type and chimaeric cDNAs were transiently transfected into COS-7 cells as described<sup>6</sup>. Forty-eight hours after transfection, competition binding experiments were done in 24-well plates (25,000-100,000 cells per well) using 20 pM [<sup>125</sup>I]-labelled CCK-8 (New England Nuclear) as the radioligand and L365260 as the competitor. After incubating the cells for 80 min at 37 °C in the presence of radioligand and competitor, monolayers were washed three times, hydrolysed in 1M NaOH, and the bound radioactivity quantified<sup>6</sup>. Competition data were analysed using computerized nonlinear curve fitting (Inplot 4.0, GraphPad, San Diego). Results are shown as means and 95% confidence intervals from at least two independent experiments. N, amino terminus; C, carboxy terminus.

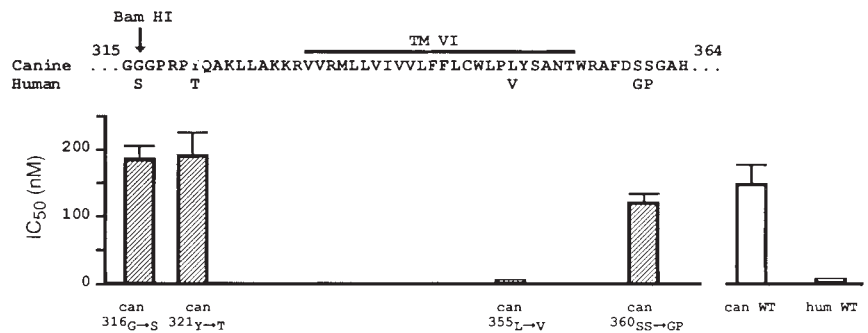
### seven-transmembrane-domain peptide hormone receptor has been identified.

To identify the functional domain underlying antagonist specificity, we have exploited the pharmacological differences in CCK-B/gastrin receptors between species. Chimaeric human/canine receptors were transiently expressed in COS-7 cells and their affinities for L365260 determined by competition binding experiments. [<sup>125</sup>I]-labelled CCK-8 was chosen as the radioligand because of its high affinity ( $K_d \sim 100$  pM) for both the human<sup>7</sup> and canine<sup>6</sup> wild-type receptors. The *BamHI* chimaera (Fig. 1) revealed that the segment of the human receptor extending from the distal end of the third intracellular loop through the carboxy terminus was sufficient to maintain high affinity for L365260. The amino-acid sequences of the corresponding region of the human and rat receptors (high affinity for L365260) (ref. 11) were compared with the canine receptor sequence (low affinity for L365260). Only five amino acids were unique to the canine receptor and were therefore candidates to explain the interspecies variation in antagonist affinity. Substitution of these five canine amino acids with the corresponding human residues resulted in four mutant canine receptors; one with two amino-acid changes (Fig. 2). Each construct was expressed in COS-7 cells and tested for L365260 affinity by [<sup>125</sup>I]-labelled CCK-8 competition binding. Replacement of Leu 355 in the canine receptor with the corresponding amino acid from the human protein, Val 319, increased the affinity for L365260 to the human wild-type value. Amino-acid substitutions in the other four positions had no significant effect; IC<sub>50</sub> (concentration giving 50% inhibition) values remained the same as

‡ To whom correspondence should be addressed.



FIG. 2 Identification of the amino-acid residue of the CCK-B/gastrin receptor that determines affinity for L365260. Amino acids (single-letter code) 315 to 364 of the canine CCK-B/gastrin receptor are compared with the corresponding human sequence; only human residues that diverge are shown. Using oligonucleotide-directed site-specific mutagenesis<sup>14,15</sup>, the nucleotides encoding the human amino acids were introduced into the same position of the canine cDNA. Mutants were confirmed by dideoxy nucleotide sequencing. Oligonucleotides used to introduce mutations included: canine (can)316G→S (5'-CTCCTGGG-C C CGGATCCGGCCCCCGGCC-3'), can321Y→T (5'-GG-CCCCGGGCCACCCAGGCCAAGCTG-3'), can355L→V (5'-CCTGTGTTGGTGGCCAGTGTATAGTGCCAACACG-3'), can360SS→GP (5'-CGTGGCGTGCCTCGACGGCCCTGGTGACACCGCGC-3'). Wild-type and mutant cDNAs were cloned into the expression vector pcDNA 1, transfected into COS-7 cells, and characterized by [<sup>125</sup>I]-labelled CCK-8 competition binding experiments with L365260 as described for Fig. 1. Results are expressed as means ± standard errors from at least two experi-



ments. Cross-hatched and open columns represent mutant and wild-type receptors, respectively. Two independently generated mutants were tested for each variant of the canine receptor. WT, wild type; can, canine; hum, human; TM, transmembrane domain.

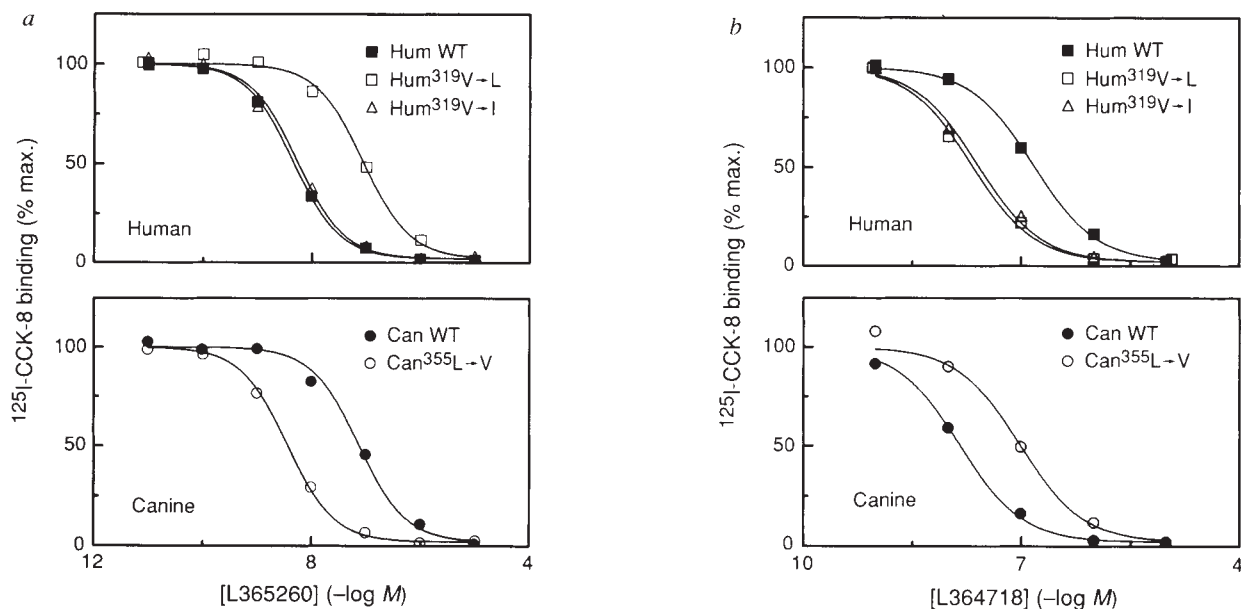


FIG. 3 Comparison of non-peptide antagonist binding to human and canine wild-type and mutant receptors. [<sup>125</sup>I]-labelled CCK-8 competition binding experiments were done with L365260 (a) and L364718 (b). Site-directed mutagenesis, transient expression of the recombinant receptors in COS-7 cells, and radioligand binding assays were done as described for Figs 1 and 2. Oligonucleotides used to generate the mutant human (hum) receptors hum319V→L and hum319V→I were 5'-CTGTGTTGGTGGCCATTATAGTG-CCAACACG-3' and 5'-TGTGTTGGTGGCAATTATAGTGCCAACA-3', respectively. All mutants were confirmed by dideoxy nucleotide sequencing over

the entire protein coding region. Data represent means from at least 4 independent competition experiments. IC<sub>50</sub> values and 95% confidence intervals for L365260 and L364718 affinities, respectively, were: humWT, 4.5 nM (4.3–4.8) and 147 nM (123–191); hum319V→L, 90 nM (54–151) and 21 nM (11–40); hum319V→I, 5.3 nM (4.7–6.0) and 26 nM (18–40); canWT, 80 nM (65–100) and 14 nM (12–17); can355L→V, 3.6 nM (3.0–4.3) and 95 nM (40–223). It was confirmed by control experiments (<sup>125</sup>I)-labelled CCK-8 competition with unlabelled CCK-8) that the radioligand affinity was essentially identical in wild-type and mutant receptors (not shown).

that of the canine wild-type receptor (Fig. 2). The converse experiment was carried out by changing the corresponding amino acid in the human receptor, Val 319, to a leucine residue (319V→L). The affinity of the human 319V→L mutant for L365260 was reduced to that of the canine wild-type (Fig. 3a), confirming the importance of the residue in this position.

The same amino acid appears to define specificity for L364718. Replacement of Leu 355 in the canine receptor with the corresponding amino acid from the human protein, Val 319, decreased the affinity for L364718 to the human wild-type value (Fig. 3b). Conversely, a mutation in the human receptor replacing Val 319 with leucine increased the binding affinity of L364718 to that of the canine wild type. It is apparent from these results that Val 319 in the human receptor and Leu 355 in the canine receptor are critical determinants of antagonist affinity for both benzodiazepine-based compounds.

When the amino-acid sequences of all known CCK/gastrin receptors were compared, the residue corresponding to human Val 319 was always found to be a branched chain aliphatic amino acid (Fig. 4). The rat and human CCK-B/gastrin receptors both have valine in this critical position. Consistent with our conclusion that this residue determines non-peptide antagonist binding, human and rat receptors have comparable affinities for L365260 as well as L364718 (refs 7, 11). A leucine in this position results in markedly altered affinities for both L365260 and L364718, as shown by the canine receptor. In the mastomys CCK-B/gastrin<sup>10</sup> and the rat CCK-A (ref. 12) receptors, an isoleucine corresponds to human Val 319. Although antagonist affinities for the mastomys receptor have not yet been reported, it is well established that the rat CCK-A receptor binds L364718 with >100-fold higher affinity, and L365260 with >100-fold lower affinity than the human CCK-B/gastrin receptor.

|               |                         |
|---------------|-------------------------|
| CCK-B/gastrin |                         |
| Human         | VVRMLLVIVLVFLCWLPVYSANT |
| Canine        | -----L-----             |
| Rat           | -----L-----V--V--       |
| Mastomys      | -----L-----I----        |
| CCK-A         |                         |
| Rat           | -I---I-----M-I-F---A    |

FIG. 4 Comparison of amino-acid sequences (single-letter code) corresponding to transmembrane domain VI of human, canine, rat, and mastomys CCK-B/gastrin and rat CCK-A receptors, as deduced from their cDNAs. Dashes indicate identity with the human receptor. Boxed amino acids correspond to human Val 319.

Therefore a mutant human CCK-B/gastrin receptor, 319V→I, was constructed to determine whether this amino-acid substitution is sufficient to confer CCK-A type antagonist selectivity. As shown in Fig. 3a and b, the isoleucine substitution increased the affinity of the human CCK-B/gastrin receptor for L364718 while leaving the affinity for L365260 essentially unchanged. The  $IC_{50}$  values of human 319V→I are considerably different from those expected for a CCK-A receptor, suggesting that other residues also contribute to the differences in antagonist affinities between the CCK-B/gastrin and the CCK-A receptors.

Our results demonstrate the role of individual amino acids in determining ligand affinities for a seven-transmembrane-domain peptide hormone receptor. One other analogous example has been reported from the biogenic amine family of receptors, in which divergence at a single amino-acid position results in significant pharmacological differences between the rodent (asparagine) and the human (threonine) 5-hydroxytryptamine (5-HT<sub>1B</sub>) receptors<sup>13</sup>. In contrast to the CCK-B/gastrin receptor, the critical amino acid in the 5-HT<sub>1B</sub> receptor is located in transmembrane domain VII rather than VI. The amino-acid difference between 5-HT<sub>1B</sub> receptor homologues results in major shifts of both agonist and antagonist affinities, whereas the species-dependent polymorphisms of the CCK-B/gastrin receptor, as described here, have significant influence on antagonist binding only.

All known CCK-B/gastrin receptors have comparable agonist affinities, despite the variability in the aliphatic amino acid residue Val (human, rat<sup>7,11</sup>); Leu (canine<sup>6</sup>), Ile (mastomys<sup>10</sup>) at the critical antagonist binding position. This contrasts with the sensitivity of non-peptide antagonist binding to changes in this single amino acid. The affinity determinants of agonists and non-peptide antagonists must therefore, at least in part, be different.

The observation that variability in the aliphatic side chain of a single amino acid in transmembrane domain VI of the CCK-B/gastrin receptor is sufficient to reverse the affinity rank order of the non-peptide receptor antagonists L365260 and L364718 could have pharmacological implications. The potential for altered ligand specificity arising from species-dependent polymorphisms will be important in drug development. The isoleucine, leucine and valine substitution mutants (Fig. 3a and b) illustrate that antagonist affinities may be dramatically altered by even a highly conservative amino-acid substitution. Receptor antagonists used for the treatment of human disease should be evaluated only in animals having receptors that are pharmacologically identical to the human homologue.

**Note added in proof:** While this manuscript was in press, Sachais *et al.* reported an additional example illustrating the molecular basis for pharmacological diversity between peptide hormone receptors from different species (*J. biol. Chem.* **268**, 2319–2323; 1993). A 90-fold affinity difference for the synthetic antagonist CP 96345 between the rat and the human substance P receptors is primarily explained by species divergence in residue 290 (TM VII) with minor contributions by differing amino acids in TM VI and in the second extracellular loop. □

Received 11 December 1992; accepted 1 February 1993.

1. Ravard, S. & Dourish, C. T. *Trends Pharmac. Sci.* **11**, 271–273 (1990).
2. Faris, P. L., Komisaruk, B. R., Watkins, L. R. & Mayer, D. J. *Science* **219**, 310–312 (1983).
3. Dourish, C. T. *et al. Eur. J. Pharmac.* **176**, 35–44 (1990).
4. Wiertelak, E. P., Maier, S. F. & Watkins, L. R. *Science* **256**, 830–833 (1992).
5. Dourish, C. T., Rycroft, W. & Iversen, S. D. *Science* **245**, 1509–1511 (1989).
6. Kopin, A. S. *et al. Proc. natn. Acad. Sci. U.S.A.* **89**, 3605–3609 (1992).
7. Lee, Y. M. *et al. J. biol. Chem.* (in the press).
8. Lotti, V. J. & Chang, R. S. *Eur. J. Pharmac.* **162**, 273–280 (1989).
9. Chang, R. S. & Lotti, V. J. *Proc. natn. Acad. Sci. U.S.A.* **83**, 4923–4926 (1986).
10. Nakata, H. *et al. Biochem. biophys. Res. Commun.* **187**, 1151–1157 (1992).
11. Wark, S. A., Piseña, J. R. & de Weerth, A. *Proc. natn. Acad. Sci. U.S.A.* **89**, 8691–8695 (1992).
12. Wark, S. A. *et al. Proc. natn. Acad. Sci. U.S.A.* **89**, 3125–3129 (1992).
13. Oksenberg, D. *et al. Nature* **360**, 161–163 (1992).
14. Kurkel, T. A. *Proc. natn. Acad. Sci. U.S.A.* **82**, 488–492 (1985).
15. Eghtedarzadeh, M. K. & Henikoff, S. *Nucleic Acids Res.* **14**, 5115 (1986).

ACKNOWLEDGEMENTS. We thank Glaxo for L365260 and L364718; A. Leiter, S. L. Lee, A. Plaut, J. Nishitani, A. Kane and K.-Fr. Sewing for discussion and for support; and L. Rogers and L. Lincicome for technical assistance. This work was supported by grants from the NIDDK.

## Amino–aromatic interaction between histidine 197 of the neurokinin-1 receptor and CP 96345

Tung Ming Fong, Margaret A. Cascieri, Hong Yu, Alka Bansal, Christopher Swain\* & Catherine D. Strader

Department of Molecular Pharmacology and Biochemistry 80M-213, Merck Research Laboratories, PO Box 2000, Rahway, New Jersey 07065, USA

\* Neuroscience Research Centre, Merck Sharp and Dohme Research Laboratories, Harlow, Essex CM20 2PT, UK

SUBSTANCE P is a peptide neurotransmitter that binds to the neurokinin-1 receptor and is involved in pain transmission and neurogenic inflammation<sup>1–4</sup>. Recently, a non-peptide substance P antagonist (CP 96345) has been shown to be effective in animal models of pain and inflammation<sup>5,6</sup>. An understanding of the molecular interactions responsible for ligand binding will be critical for the development of more specific and selective antagonists for the neurokinin-1 receptor and for the discovery of new antagonists for related G-protein-coupled receptors. Here we report that histidine at position 197 in the fifth transmembrane helix of the human neurokinin-1 receptor binds specifically to CP 96345 but not to peptide agonists. Substitution of His 197 by different amino acids and analysis of structural analogues of antagonists suggest that His 197 is involved in an amino–aromatic interaction with the benzhydryl moiety of CP 96345.

The neurokinin-1 receptor (NK-1R) is a G-protein-coupled receptor characterized by seven putative transmembrane helices<sup>7–9</sup>. Because of the topological similarity among G-protein-coupled receptors and the observation that  $\beta$ -adrenergic receptors and rhodopsin bind small molecule ligands within their transmembrane domains<sup>10–12</sup>, we reasoned that the transmembrane domain of the NK-1R may be involved in the binding of CP 96345. We applied Fourier analysis of the putative transmembrane region to identify residues likely to be located at the centre of the human NK-1R near the top of the transmembrane domain<sup>13</sup>. Among those interior residues, several were selected for site-directed mutagenesis on the basis of side chains that contain functional groups capable of forming hydrogen bonds with ligands. Substituting four such residues by alanine (T40A, H95A, T201A, Q284A; Fig. 1) did not affect the binding affinity of substance P (SP) or CP 96345 (Table 1), suggesting that these residues are not involved in ligand binding. In contrast, substitution of His 197 in helix 5 by alanine (H197A) resulted in a 30-fold reduction in CP 96345 affinity without affecting substance



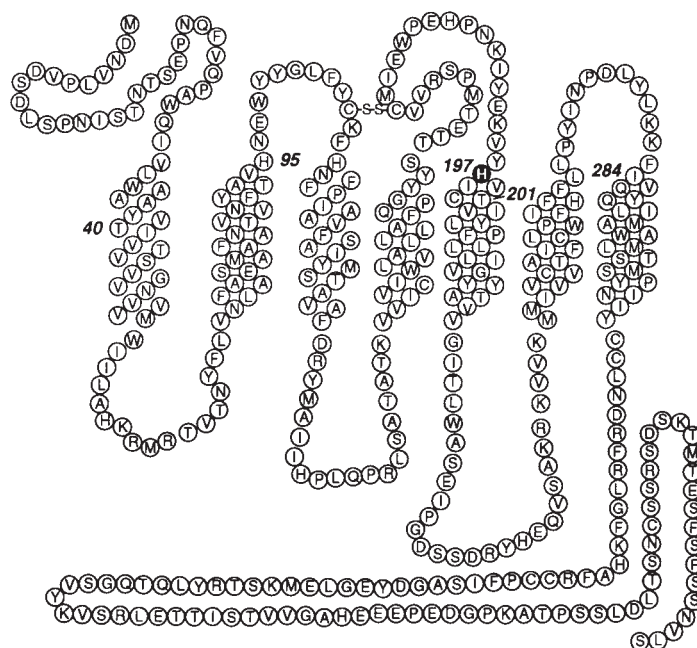


FIG. 1 Schematic model of the human NK-1R showing putative transmembrane helices and the relative locations of the residues mutated. His 197 residue is highlighted.

P affinity (Table 1), suggesting that His 197 is specifically involved in antagonist binding. In order to probe the nature of the interaction between His 197 and CP 96345, several different substitutions were introduced at this position. The H197Q mutant bound CP 96345 with the same affinity as the wild-type receptor (Table 1), indicating that the aromatic characteristics of His 197 are not required for the interaction with the antagonist. Substitution by phenylalanine (H197F) caused only a three-fold reduction in CP 96345 affinity, suggesting that a hydrogen-bonding interaction is not critical. The absence of hydrogen bonds between His 197 and CP 96345 is also consistent with the results of the H197Y and H197S mutants, both of which exhibited about 15-fold reduction in CP 96345 affinity compared to the wild-type NK-1R. Replacement of His 197 by lysine

(H197K) resulted in a 500-fold reduction in CP 96345 affinity. Therefore, only glutamine and, to a lesser extent, phenylalanine could substitute directly for His 197 in the interaction with CP 96345 (Table 1).

The interaction between CP 96345 and His 197 is specific because the binding affinity for the perhydroisoindole antagonist RP 67580 (ref. 14) was not affected by the H197A mutation (Table 1). This is consistent with previous data indicating that different antagonists bind to overlapping but non-identical sites<sup>15</sup>. Most of the point mutants described here retained normal binding affinity for substance P (Table 1), indicating that the gross conformation of the NK-1R is not perturbed by the substitution of His 197. One mutant (H197K) showed a 5-fold reduction in affinity for substance P but not in the affinities of other neurokinin peptides (data not shown), which may indicate a perturbation of the substance P-receptor interaction as a result of the longer side chain and/or the positive charge of Lys 197.

To determine which functional group in CP 96345 interacts with His 197, we analysed a series of CP 96345 analogues in which the major substituents were altered. If a functional group on CP 96345 interacts with His 197, removing that functional group should result in a dissociation constant for the H197A mutant equal to that for the wild type because the His 197 specific interaction is no longer present. On the other hand, removal of a functional group interacting with a residue other than His 197 should lead to a decrease in affinity for the H197A mutant compared with the wild type, analogous to what is observed for the parent compound CP 96345. Thus replacement of the quinuclidine nitrogen of CP 96345 by carbon yielded L706164, with an affinity for the human NK-1R substantially lower than CP 96345 (Table 2). This compound showed further reduced affinities for the series of substitution mutants with a pattern similar to CP 96345; that is, glutamine was a perfect substitution for His 197, whereas serine, alanine and lysine were not (Fig. 2a). The additive effects of substituting His 197 on the NK-1R and replacing the quinuclidine nitrogen on CP 96345 indicate that this nitrogen interacts with a residue other than His 197. Similarly, modification of the secondary amine (L703605) or the *N*-benzyl substituent (L705084) resulted in the same pattern of relative affinities for the mutant receptors as CP 96345 (Fig. 2a), suggesting that neither the secondary amine nor the substituted benzyl moiety interacts with His 197.

TABLE 1 Binding affinities for wild-type and mutant receptors

| Receptor  | IC <sub>50</sub> (nM) |                |              |
|-----------|-----------------------|----------------|--------------|
|           | Substance P           | CP 96345       | RP 67580     |
| Wild type | 0.63 ± 0.15 (7)       | 0.45 ± 0.1 (4) | 18 ± 3 (4)   |
| T40A      | 1 (2)                 | 0.7 (2)        | 30 (2)       |
| H95A      | 1 (2)                 | 0.9 (2)        | 13 (2)       |
| T201A     | 1.1 ± 0.4 (2)         | 0.7 ± 0.1 (2)  | 32 ± 2 (2)   |
| Q284A     | 0.6 ± 0.1 (2)         | 0.3 (2)        | 22 ± 9 (2)   |
| H197A     | 0.68 ± 0.1 (2)        | 17 ± 3 (2)     | 22 ± 3 (2)   |
| H197Q     | 1.1 ± 0.1 (2)         | 0.4 ± 0.1 (2)  | 16 ± 4 (2)   |
| H197F     | 0.7 (2)               | 1.2 ± 0.1 (2)  | 49 ± 11 (2)  |
| H197Y     | 1.1 ± 0.4 (2)         | 5.2 ± 0.6 (2)  | 48 ± 19 (2)  |
| H197S     | 1.2 ± 0.2 (2)         | 6.8 ± 0.1 (2)  | 30 ± 7 (2)   |
| H197K     | 3.0 ± 0.6 (4)         | 265 ± 60 (4)   | 264 ± 36 (3) |

Point mutations were created by oligonucleotide-mediated site-directed mutagenesis and confirmed by DNA sequencing. Binding was assayed using intact COS cells expressing the cloned human NK-1R or its mutants<sup>9</sup>. The data were fitted to the equation  $(c.p.m.(L) - c.p.m.(1 \mu M SP)) / (c.p.m.(0) - c.p.m.(1 \mu M SP)) = IC_{50} / (L + IC_{50})$ , in which c.p.m.(L) and c.p.m.(0) represent the amount of bound <sup>125</sup>I-labelled substance P in the presence or in the absence of unlabelled ligand, respectively; L represents the concentration of unlabelled ligand, and IC<sub>50</sub> represents the concentration of unlabelled ligand that causes 50% inhibition of specifically bound <sup>125</sup>I-labelled substance P. Either Bolton-Hunter <sup>125</sup>I-labelled substance P or [<sup>125</sup>I] Tyr8-substance P were used in the assay; both radioligands gave identical IC<sub>50</sub> values. The means ± s.d. from independent experiments are listed.

TABLE 2 Binding affinities of quinclidine analogues for the human NK-1R and His 197 mutants

| Compound | X | R <sub>1</sub> | R <sub>2</sub> | R <sub>3</sub>  | R <sub>4</sub>   | IC <sub>50</sub> (nM) |                 |                   |                    |                    |                    |
|----------|---|----------------|----------------|-----------------|------------------|-----------------------|-----------------|-------------------|--------------------|--------------------|--------------------|
|          |   |                |                |                 |                  | WT                    | H197Q           | H197F             | H197S              | H197A              | H197K              |
| CP 96345 | N | Ph             | Ph             | H               | OCH <sub>3</sub> | 0.45 ± 0.1 (4)        | 0.4 ± 0.1 (2)   | 1.2 ± 0.1 (2)     | 6.8 ± 0.1 (2)      | 17 ± 2 (2)         | 265 ± 60 (4)       |
| L706164  | C | Ph             | Ph             | H               | OCH <sub>3</sub> | 1,000 ± 350 (4)       | 1,100 ± 900 (2) | 5,200 ± 2,200 (2) | 9,000 ± 1,000 (2)  | 12,000 ± 3,500 (2) | >20,000 (2)        |
| L705084  | N | Ph             | Ph             | CH <sub>3</sub> | OCH <sub>3</sub> | 55 ± 15 (9)           | 63 ± 50 (2)     | 41 ± 19 (2)       | 1,600 (2)          | 2,500 (2)          | >20,000 (2)        |
| L703605  | N | Ph             | Ph             | H               | H                | 118 ± 81 (2)          | 101 ± 78 (2)    | 370 ± 25 (3)      | 1,221 ± 480 (2)    | 2,050 ± 250 (2)    | 10,000 ± 500 (2)   |
| L703766  | N | Ph             | H              | H               | OCH <sub>3</sub> | 2,350 ± 1,000 (4)     | 1,700 ± 300 (2) | 3,500 ± 2,500 (2) | 2,000 ± 750 (2)    | 2,800 ± 1,300 (2)  | 4,800 ± 800 (2)    |
| L703774  | N | M              | M              | H               | OCH <sub>3</sub> | 206 ± 64 (3)          | 375 ± 75 (2)    | 1,850 ± 150 (2)   | 8,500 ± 6,000 (2)  | 9,000 ± 3,900 (3)  | 12,000 ± 3,700 (3) |
| L705081  | N | M              | H              | H               | OCH <sub>3</sub> | 6,800 ± 440 (3)       | 7,000 (2)       | 3,500 ± 500 (2)   | 10,000 ± 6,000 (2) | 8,200 ± 4,000 (3)  | 6,200 ± 2,300 (3)  |

Abbreviations: WT, wild type; Ph, phenyl; M, *p*-methoxyphenyl. All compounds are 2, 3-*cis* racemates, and were synthesized as described<sup>25</sup>. The binding assay was as described in Table 1 legend, except in the case of L706164, L705084, L703766, L703774 and L705081, for which binding was assayed using 0.2 nM [<sup>125</sup>I]Tyr8-substance P in the presence of 0.05 mg ml<sup>-1</sup> chymostatin, 0.002 mg ml<sup>-1</sup> pepstatin and 0.1 mM PMSF instead of bacitracin and phosphoramidon<sup>9</sup>. The means ± s.d. from independent experiments are listed.

L703766 is a CP 96345 analogue in which the benzhydryl at C2 is replaced by a benzyl group. The affinities of L703766 for all mutants were very similar to that for the wild-type receptor (Table 2). The observation that the affinity of L703766 is insensitive to the chemical nature of the side chain at position 197,

whereas CP 96345 binding is substantially affected by substitution at this position (Fig. 2a) strongly suggests that His 197 interacts directly with the benzhydryl moiety of CP 96345. This suggestion is supported by analysis of a related compound, L703774, in which *p*-methoxy substituents are added to the benzhydryl moiety. Again, removal of one of the aromatic rings of the substituted benzhydryl (L705081) results in a loss of sensitivity to the chemical nature of the residue at position 197 (Table 2, and Fig. 2a). It should be noted that the affinity of L703766 for the wild-type receptor was substantially lower than that of CP 96345. Such a decrease is greater than expected for the loss of a single interaction between CP 96345 and the NK-1R (for example, 30-fold reduction of CP 96345 affinity in the H197A mutant), suggesting that other amino-acid residues may also participate in interactions with the benzhydryl of CP 96345. The increased conformational flexibility of L703766 may also explain the weakened interaction with the receptor.

The observation that His 197 can be perfectly substituted by glutamine suggests that the imidazole of His 197 may be involved in an amino-aromatic interaction<sup>16,17</sup> with the benzhydryl of CP 96345 (Fig. 2b). According to this model, the δ(+) amino group of histidine or glutamine would form an electrostatic interaction with the δ(−) π-electrons of the benzhydryl group of the antagonist<sup>18–20</sup>. In the case of the H197F mutant, an edge-to-face aromatic-aromatic interaction<sup>18,21,22</sup> between the δ(+) hydrogen of the phenylalanine and the δ(−) π-electrons of the benzhydryl group of the antagonist would be similar to an amino-aromatic interaction in its electrostatic origin. Theoretical calculations have suggested that an electrostatic amino-aromatic interaction can contribute ~3 kcal mol<sup>-1</sup> of binding enthalpy<sup>23,24</sup>, which is consistent with the 30-fold reduction in CP 96345 affinity for the H197A mutant (Fig. 2a). In contrast, the longer side chain of lysine in the H197K mutant may create an unfavourable positioning for CP 96345, whereas an unfavourable contribution from the δ(−) oxygen of serine or tyrosine in the H197S or H197Y mutants might preclude a favourable interaction with the benzhydryl group of CP 96345.

The conclusion that His 197 interacts directly with the benzhydryl of CP 96345 suggests that other nearby residues within the transmembrane domain should also interact with the antagonist<sup>15</sup>. The identification of a specific molecular interaction between CP 96345 and the NK-1R should provide a rational basis for the development of more potent and selective antagonists. To our knowledge, these data are the first evidence for a specific interaction between a small molecule antagonist and the transmembrane domain of a peptide receptor. Together with previous data showing a similar location for the small molecule binding sites in adrenergic receptors<sup>11</sup>, these results suggest that there is a general transmembrane binding site for small

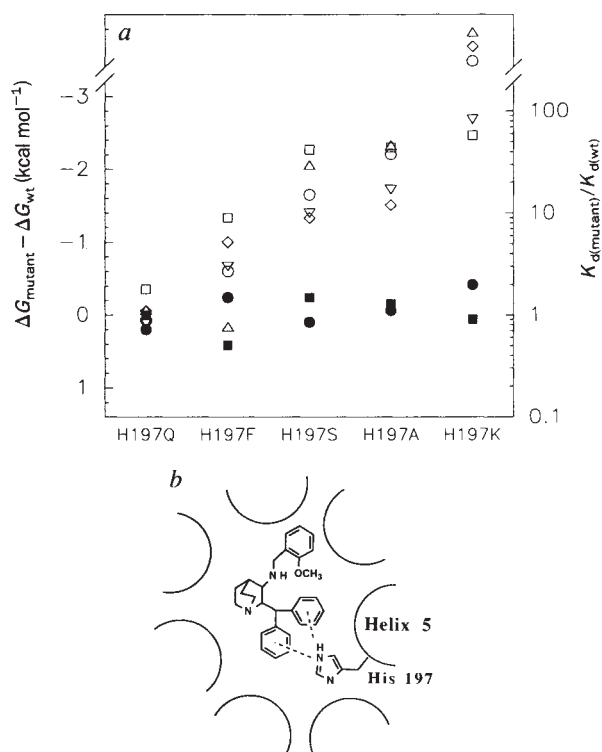


FIG. 2 a, Difference in the calculated free energies of dissociation between a mutant and the wild-type receptor for various compounds. The free energy of dissociation is calculated according to the equation  $\Delta G = -RT \ln(K_d)$ , in which the dissociation constant for a compound ( $K_d$ ) is calculated as  $IC_{50}/(1 + [^{125}I\text{-SP}]/K_{d(\text{SP})})$ , and the dissociation constant for substance P ( $K_{d(\text{SP})}$ ) is calculated as  $IC_{50(\text{SP})} - [^{125}I\text{-SP}]$ . Right axis shows the ratio of dissociation constants for a mutant and wild-type receptor. Left axis shows the difference in the calculated free energies of dissociation ( $-RT \ln(K_{d(\text{mutant})}/K_{d(\text{wt})})$ ). Each symbol represents one compound: CP 96345 (○), L703766 (●), L703774 (□), L705081 (■), L705084 (△), L703605 (▽) and L706164 (◇). Data points outside the vertical axis break represent a mutant-to-wild type ratio of higher than 100, or a  $K_{d(\text{mutant})}$  value larger than 20 μM. b, Schematic drawing of the proposed amino-aromatic interaction between His 197 and CP 96345 in the transmembrane core of the NK1R. Each arc represents a transmembrane helix.



molecules in G-protein-coupled receptors, and that non-peptide antagonists can be developed for the analogous domain in other peptide receptors. □

Received 7 January; accepted 11 February 1993.

- Iversen, L. L., Watling, K. J., McKnight, A. C., Williams, B. J. & Lee, C. M. *Top. Med. Chem.* **65**, 1-9 (1987).
- Regoli, D., Drapeau, G., Dion, S. & D'Orléans-Juste, P. *Pharmacology* **38**, 1-15 (1989).
- Helke, C. J., Krause, J. E., Mantyh, P. W., Couture, R. & Bannon, M. J. *FASEB J.* **4**, 1606-1615 (1990).
- Nakanishi, S. *A. Rev. Neurosci.* **14**, 123-136 (1991).
- Snider, R. M. *et al. Science* **251**, 435-437 (1991).
- DeKoninck, Y. & Henry, J. L. *Proc. natn. Acad. Sci. U.S.A.* **88**, 11344-11348 (1991).
- Yokota, Y. *et al. J. Biol. Chem.* **264**, 17649-17652 (1989).
- Hershey, A. D. & Krause, J. E. *Science* **247**, 958-962 (1990).
- Fong, T. M., Anderson, S. A., Yu, H., Huang, R.-R. C. & Strader, C. D. *Molec. Pharmac.* **41**, 24-30 (1992).
- Khorana, H. G. *J. Biol. Chem.* **267**, 1-4 (1992).
- Strader, C. D., Sigal, I. S. & Dixon, R. A. F. *FASEB J.* **3**, 1825-1832 (1989).
- Ward, W. H. J., Timms, D. & Fersht, A. R. *Trends Pharmac. Sci.* **11**, 280-284 (1990).
- Rees, D. C., DeAntonio, L. & Eisenberg, D. *Science* **245**, 510-513 (1989).
- Garrett, C. *et al. Proc. natn. Acad. Sci. U.S.A.* **88**, 10208-10212 (1991).
- Fong, T. M., Yu, H. & Strader, C. D. *J. Biol. Chem.* **267**, 25668-25671 (1992).
- Burley, S. K. & Petsko, G. A. *Adv. Protein Chem.* **39**, 125-189 (1988).
- Singh, J. & Thornton, J. M. *J. molec. Biol.* **211**, 595-615 (1990).
- Perutz, M. F., Fermi, G., Abraham, D. J., Poyart, C. & Bursaux, E. *J. Am. chem. Soc.* **108**, 1064-1078 (1986).
- Waksman, G. *et al. Nature* **358**, 646-653 (1992).
- Dougherty, D. A. & Stauffer, D. A. *Science* **250**, 1558-1560 (1990).
- Burley, S. K. & Petsko, G. A. *Science* **229**, 23-28 (1985).
- Singh, J. & Thornton, J. M. *FEBS Lett.* **191**, 1-6 (1985).
- Deakne, C. A. & Meot-Ner (Mautner), M. *J. Am. chem. Soc.* **107**, 474-479 (1985).
- Levitt, M. & Perutz, M. F. *J. molec. Biol.* **201**, 751-754 (1988).
- Lowe, J. A. *et al. J. med. Chem.* **35**, 2591-2600 (1992).

ACKNOWLEDGEMENTS. We thank T. Ladduwhetty, M. Vannell, V. Sabin, E. Seward and B. Williams of the Neuroscience Research Center for synthesizing CP 96345 and its analogues, R.-R.C. Huang and L.-L. Shiao for help with binding assays, and D. Underwood for discussing molecular modelling and for reviewing the manuscript.

## SNAP family of NSF attachment proteins includes a brain-specific isoform

Sidney W. Whiteheart\*, Irene C. Griff†, Michael Brunner\*, Douglas O. Clary\*, Thomas Mayer\*, Susan A. Buhrow\* & James E. Rothman\*§

\* Rockefeller Research Laboratory, Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, New York, New York 10021, USA

**THE soluble NSF attachment proteins (SNAPs) enable *N*-ethylmaleimide-sensitive fusion protein (NSF) to bind to target membranes<sup>1-3</sup>. Here we report the cloning and sequencing of complementary DNAs encoding  $\alpha$ -,  $\beta$ - and  $\gamma$ -SNAPs. Two of these proteins,  $\alpha$  and  $\gamma$ , are found in a wide range of tissues, and act synergistically in intra-Golgi transport. The third,  $\beta$ , is a brain-specific isoform of  $\alpha$ -SNAP. Thus, NSF and SNAPs appear to be general components of the intracellular membrane fusion apparatus, and their action at specific sites of fusion must be controlled by SNAP receptors particular to the membranes being fused, as described in the accompanying article<sup>4</sup>.**

SNAPs purified from bovine brain<sup>3</sup> were digested by protease to give peptides whose sequences were used to generate probes for the screening of a bovine brain cDNA library (see legend to Fig. 1). The cDNA clones for  $\alpha$ -,  $\beta$ - and  $\gamma$ -SNAP contain open reading frames of 885, 894 and 984 nucleotides, encoding proteins with predicted  $M_r$ s of 33K, 34K and 36K, respectively. Figure 1 shows the amino-acid sequence of the three SNAPs compared to the proposed yeast homologue of  $\alpha$ -SNAP, SEC17 (refs 2, 5).  $\alpha$ - and  $\beta$ -SNAP are 83% identical to each other and 34% and 33% identical to SEC17, respectively.  $\gamma$ -SNAP is less

|            |     |             |            |            |                    |
|------------|-----|-------------|------------|------------|--------------------|
| Alpha SNAP | 1   | MDNSG       | KEAEAMALLA | EAEKRVKNSQ | SFFSGLFEGGS        |
| Beta SNAP  |     | MDNAG       | KEREAVOLMA | EAEKRVKASH | SELRLGFGN          |
| Gamma SNAP |     | MDGAGASAST  | SGPRREMAAQ | KINGLEHLA  | KAEK.....YLKTFGLKW |
| Sec17p     |     |             |            | MSPDPEVLEK | EAEKRGVPES         |
|            |     |             |            |            | GEMKLFSSSD         |
| Alpha SNAP | 51  | S.KIEEACEI  | YAAANMFM   | AKNWSAAGSA | ECORAHVHLQ         |
| Beta SNAP  |     | T.RIEEACEM  | YTRAANMFM  | AKNWSAAGNA | ECORAKLHMQ         |
| Gamma SNAP |     | KPDYDSKASE  | YGKAAVAFKN | AKOFEQACDA | CLKKAVAHEN         |
| Sec17p     |     | SYKFEENADL  | CVQAAITYRL | RKEINLAGDS | ELKAAIDYQKK        |
|            |     |             |            |            | AGNEDAGNT          |
| Alpha SNAP | 101 | FVDAGNAFKK  | .ADPQEAINE | EMRAEITYD  | MSRFTIAAKH         |
| Beta SNAP  |     | FVDAGNAFKK  | .ADPQEAINE | EMRAEITYD  | MSRFTIAAKH         |
| Gamma SNAP |     | YEQAGHMLKE  | MOKLPERVQL | IEKASHMYLE | NGPTDTAAMA         |
| Sec17p     |     | YERAYKCFKS  | GGNSVNAVDS | LENFIQIFTH | RGQPRGANF          |
|            |     |             |            |            | KFELGETLEN         |
| Alpha SNAP | 151 | ELVDIEKATA  | HYEQSADYIK | GEESNSPANK | CLLRVAGYAA         |
| Beta SNAP  |     | ELVDIEKATA  | HYEQSADYIK | GEESNSPANK | CLLRVAGYAA         |
| Gamma SNAP |     | ..VDEKAVQ   | LYQGTANVE  | NEERLQAVE  | LAKASRLLV          |
| Sec17p     |     | DLHDYAKAID  | CYELAGEWYA | QDOVALNSK  | CFIKCADLKA         |
|            |     |             |            |            | LDGQYIEASD         |
| Alpha SNAP | 201 | TYEQVGTNAM  | DSPELKYSAK | DYFFKAALCH | FCIDMLNAKL         |
| Beta SNAP  |     | TYEQVGTNAM  | DSPELKYSAK | DYFFKAALCH | FCIDMLNAKL         |
| Gamma SNAP |     | SIQKEKNITYK | EIENYPTCYK | K.TIAQVIVH | LHNDYVYAR          |
| Sec17p     |     | TYSKLIKSSK  | GNRLSQWISK | DYFLKGLQ   | LAATDAVAAR         |
|            |     |             |            |            | TLQEQGSEDE         |
| Alpha SNAP | 251 | AFSDSRECKR  | IKKLELAHEE | QNVDSYTAIV | KEY.....D          |
| Beta SNAP  |     | AFSDSRECKL  | IKKLELAHEE | QNVDSYTAIV | KEY.....D          |
| Gamma SNAP |     | GENGSEDCAA  | LEQLLEGVDQ | QDQDQVAVCV | NSPLFKYMDN         |
| Sec17p     |     | NEADSRSENF  | LKSLIDAVNE | GDSEQLSEHC | KEF.....D          |
|            |     |             |            |            | NFMRLDKWKI         |
| Alpha SNAP | 301 | THMLRIKKTIT | QGDDE...DL | R          | 338                |
| Beta SNAP  |     | THMLRIKKTIT | QGDDEGGDGL | K          |                    |
| Gamma SNAP |     | VPGGGVKKKA  | AAPPAQKPEG | TAAFAAEEBE | DEYAGGIC           |
| Sec17p     |     | THMLRIKKTIT | QGDDEDL    |            |                    |

FIG. 1 Comparison of the protein sequences of the members of the SNAP family (single-letter amino-acid code). The ORFs of the corresponding cDNA clones were translated and compared by best-fit alignment<sup>14</sup>. The underlined sequences represent the chemical sequences of peptides produced by protease digestion of the intact proteins (trypsin digestion for  $\alpha$ -SNAP and Lys-C digestion for  $\beta$  and  $\gamma$ -snaps). Shading highlights the similarities between sequences.

**METHODS.** Cloning of the cDNA for  $\alpha$ -SNAP:  $\alpha$ -SNAP was digested by TPCK (tosylphenylalanine-chloromethylketone)-treated trypsin and the resulting peptides were fractionated by reversed-phase HPLC (Vydac C18 column, 0.1% TFA, 0-60% acetonitrile). Sequences of three of these peptides were compared to the Sec17p sequence<sup>5</sup> and polymerase chain reaction (PCR) primers were designed on the basis of the alignment of the three peptides. These primers (peptide 1, GCNGARATYTAAGARACNGA-OH; peptide 2, ACRTTYTGTCYTCRTGNGC-OH; peptide 3, CATNGTNGTARCCAYTG-OH) were used to a two-stage PCR with bovine brain first-strand cDNA as template, to generate probes for screening a  $\lambda$ gt10 cDNA library from bovine brain (gift from P. Seeburg, ZMBH, Heidelberg). Five clones were isolated from  $5 \times 10^5$  independent phage plaques; of these, two represented a full-length clone. ( $r=A, G; y=T, C; n=A, T, C, G$ ). Cloning of the cDNA for  $\beta$ -SNAP: Comparison of the six  $\beta$ -SNAP peptides with the predicted amino-acid sequence of the I47 partial cDNA identified from searching the database with the sequence of  $\alpha$ -SNAP showed that I47 was probably the mouse  $\beta$ -SNAP gene, as each of the amino-acid differences in the sequenced peptides between  $\alpha$ - and  $\beta$ -SNAP were found in this I47 sequence. The missing 5' end of this mouse cDNA was amplified from a  $\lambda$ ZAP mouse brain cDNA library (Stratagene) using the T7 sequencing primer and a primer specific to nucleotides 328 to 354 of clone I47 (GCCACCTTCAG-CAGACACTTGTTGG-OH). Direct sequencing of this 420 bp PCR product confirmed that it encoded the predicted 5' end of the mouse  $\beta$ -SNAP. Moreover, its nucleotide sequence is only 79% identical to the nucleotide sequence of  $\alpha$ -SNAP, so this short DNA fragment was used to screen specifically for the bovine  $\beta$ -SNAP cDNA using standard conditions.  $1.0 \times 10^6$  independent phage plaques from a bovine brain  $\lambda$ gt-10 library (Clontech) were screened, yielding 5 individual isolates, 1 of which was full-length. Cloning of the cDNA for  $\gamma$ -SNAP: Primers based on the sequences of the peptide MMYL...DPEK (5' primer ATGATGTAYTNGARAAYGG-OH, 3' primer YTTYTCNGGRTCNACRTT-OH) were used to generate suitable probes by PCR using a  $\lambda$ ZAP bovine brain cDNA library (Stratagene) as template. The T3 and T7 sequencing primers of the  $\lambda$ ZAP vector and the peptide-specific 5' primer were used to amplify. The specific products generated were fractionated by agarose gel electrophoresis and individual products were reamplified by PCR using the peptide-specific 5' and 3' primers. A fragment of the exact predicted size (90 bp) was subcloned, sequenced and used as an RNA probe with hybridization and washing conditions described (Promega).  $1.5 \times 10^6$  independent phage plaques from a bovine brain  $\lambda$ gt-10 library (Clontech) were screened yielding 5 individual isolates, one of which was full-length.

§ Present addresses: Institut für Physiologische Chemie, Universität München, Goethestrasse 33, D-8000 München, Germany (M.B.); HHMI University of California San Francisco, San Francisco, California 94143-0724, USA (D.O.C.).

related but may have similar overall structure, because it is 25% identical to  $\alpha$  and 23% identical to  $\beta$  and to SEC17. Only one related sequence was found in the Genbank, SwissProt and PIR databases. This was the mouse clone 147<sup>6</sup>, which proved to be mouse  $\beta$ -SNAP. Analysis of the SNAP sequences (ProSearch Software 1.1) identified several potential phosphorylation sites.  $\alpha$ -SNAP contains three potential protein kinase C sites (S-4, 35, 201), and seven putative casein kinase II sites (S-4, 36, 77, 126, 201, 258, 270).  $\beta$ -SNAP contains five putative casein kinase II sites (T-36, S-77, S-84, T-186 and S-199).  $\gamma$ -SNAP contains only one consensus site for tyrosine phosphorylation (Y-292).  $\alpha$ -SNAP contains three predicted coiled-coil domains (residues 5–28, 127–150, 166–188);  $\beta$ -SNAP contains two predicted coiled-coil domains (residues 127–150 and 166–188);  $\gamma$ -SNAP contains only one such region (185–205)<sup>7</sup>;  $\gamma$ -SNAP also has a proline-rich region at its C terminus (312–338).

Genomic Southern blot analysis indicates that each protein is encoded by a single gene (data not shown).  $\alpha$ - and  $\gamma$ -SNAP messenger RNAs are expressed in many tissues, but  $\beta$ -SNAP mRNA seems to be restricted to brain (Fig. 2). The  $\alpha$ -SNAP probe hybridized to two messages in each tissue examined, a major species of 1.8 kilobases (kb) matching the cDNA clone and a minor species of 3.1 kb. The  $\alpha$ -SNAP probe also hybridized to a brain-specific message of 5.6 kb detected with the  $\beta$ -SNAP-specific probe. This cross-hybridization is expected because the coding regions of  $\alpha$ - and  $\beta$ -SNAP are 76% identical at the nucleotide level.  $\gamma$ -SNAP probe detected a 5.3-kb mRNA in all tissues examined. The previous description of 147, now known to be  $\beta$ -SNAP, indicated that its expression was restricted to specific regions of the brain<sup>6</sup>.  $\alpha$ -SNAP mRNA is also more highly expressed in brain than in any other tissue examined. This was confirmed at the protein level by western blotting of extracts from various tissues (data not shown).

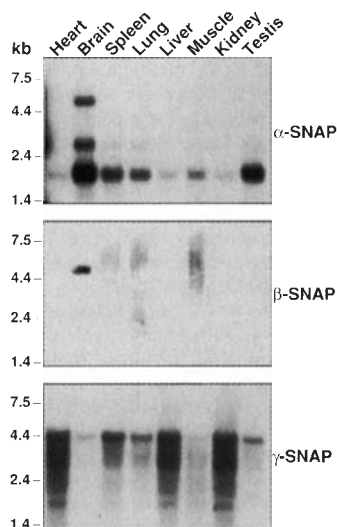


FIG. 2 Expression of  $\alpha$ -,  $\beta$ - and  $\gamma$ -SNAPs in mouse tissues. A northern blot of mouse RNA from assorted tissues (heart, brain, spleen, lung, liver, skeletal muscle, kidney, testis) was repeatedly probed with the  $\alpha$ -SNAP cDNA, the 5' end of the  $\beta$ -SNAP cDNA (described in Fig. 1) and the  $\gamma$ -SNAP cDNA.  $\alpha$ -SNAP and  $\gamma$ -SNAP gene expression is apparent in all tissues, whereas  $\beta$ -SNAP gene expression is restricted to the brain. The full coding regions of  $\alpha$ -SNAP and  $\gamma$ -SNAP as well as the 420 bp fragment of the mouse  $\beta$ -SNAP gene were used individually to probe the northern blot. These DNA fragments were isolated by agarose gel electrophoresis and then labelled with [ $\alpha$ -<sup>32</sup>P]dCTP using the random hexamer method (Boehringer-Mannheim). The northern blot was purchased from Clontech and their methods for treatment were followed.  $\alpha$ -SNAP and  $\gamma$ -SNAP probes were hybridized in 50% formamide at 42 °C for 24 h. Washes were as described by manufacturer. The  $\beta$ -SNAP probe was hybridized under aqueous conditions at 55 °C for 24 h. Hybridization of an actin probe (Clontech) demonstrated that equal amounts of actin message are present in each lane (data not shown).

To establish that the cDNAs isolated encoded active SNAPs, the cDNAs encoding  $\alpha$ - and  $\gamma$ -SNAPs were cloned into a pQE-9 vector (Qiagen) to produce proteins containing six histidine residues at their N termini. These proteins were expressed at high levels in *Escherichia coli* and purified by affinity chromatography on nickel-nitrilotriacetic acid (Ni-NTA)-agarose (Fig. 3a). Figure 3b demonstrates that the recombinant proteins were active in the intra-Golgi cell-free transport assay which they saturate at concentrations comparable to those reported for the wild-type proteins (40 ng per 50  $\mu$ l assay).  $\alpha$ - and  $\gamma$ -SNAP act synergistically in this assay when NSF is present at subsaturating concentrations (<2.5 ng per 50  $\mu$ l assay; Fig. 3b). When NSF is in excess, transport in the presence of both SNAPs is equal to the sum of the transport when each is used alone (data not shown). This synergy was not seen in earlier

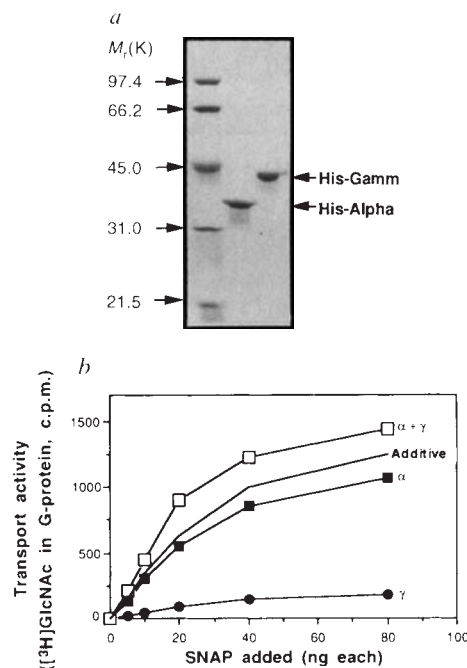


FIG. 3 Expression of recombinant  $\alpha$ - and  $\gamma$ -SNAP in *E. coli*. SNAPs containing a six-histidine extension at their N terminus were expressed in *E. coli* and purified by affinity chromatography on Ni-NTA-agarose. The pure proteins (a) were titrated into a standard transport assay using salt-washed donor and acceptor membranes<sup>3</sup> and a subsaturating amount of His<sub>6</sub>-NSF (0.2 ng per 50  $\mu$ l Assay) (b). The curve labelled 'additive' denotes the summation of VSV G-protein transport when each SNAP was used alone.

METHODS. The coding regions for each of the proteins (NSF,  $\alpha$ - and  $\gamma$ -SNAP) were engineered by PCR to contain a *Bam*HI site immediately 5' of the first ATG. The sequences of the resulting constructs were confirmed by dideoxynucleotide sequencing and then cloned into the pQE-9 vector (Qiagen). The plasmid constructs were transformed into *E. coli* (XL-1Blue, Stratagene) and the resulting cells were grown to a density of  $A_{600}=0.8$  in Supermedia (Qiagen) with 100  $\mu$ g ml<sup>-1</sup> ampicillin, 50  $\mu$ g ml<sup>-1</sup> tetracycline and were induced for 4 h with 1 mM isopropyl-thio- $\beta$ -D-galactoside. The SNAP-expressing cells were collected and washed in buffer A (25 mM Tris-HCl, pH 7.8, 75 mM KCl, 2 mM mercaptoethanol, 1 mM PMSF) disrupted in a French press, and the suspension was clarified by centrifugation at 100,000g for 1 h. The supernatant was passed over a Ni-NTA-agarose column (5 ml bed volume) and the His<sub>6</sub>-SNAPs were eluted with a 40 ml, 20–500 mM imidazole gradient in buffer A. His<sub>6</sub>- $\gamma$ -SNAP required an additional purification step on Mono Q (Pharmacia) in Tris buffer as described<sup>3</sup>. The protein concentrations were determined by the Bradford protein assay (BioRad) using ovalbumin as standard. After pooling, pure protein fractions were stored as aliquots with 10% glycerol at –80 °C. Cells expressing His<sub>6</sub>-NSF were disrupted in 100 mM HEPES/KOH, pH 7, 500 mM KCl, 5 mM MgCl<sub>2</sub>, 2 mM mercaptoethanol, 1 mM PMSF, 0.5 mM ATP. The Ni-NTA-agarose column was eluted with a 50–500 mM imidazole gradient in 20 mM HEPES/KOH, pH 7.0, 200 mM KCl, 2 mM mercaptoethanol, 0.5 mM ATP and 10% glycerol. The activity of the His<sub>6</sub>-NSF was tested using *N*-ethylmaleimide-treated donor and acceptor membranes as described<sup>10</sup>.



experiments owing to the saturating amounts of NSF used<sup>2,3</sup>. Its existence supports earlier findings that  $\gamma$ -SNAP doubles the assembly of  $\alpha$ -SNAP into 20S fusion particles<sup>8</sup> and that  $\alpha$ - and  $\gamma$ -SNAP are both needed for optimal NSF binding to Golgi membranes<sup>9</sup>. Although  $\gamma$ -SNAP is apparently not essential for cell-free transport<sup>2,3</sup> or for the assembly of 20S fusion particles<sup>8</sup> (S.W.W., unpublished results), it does appear to stabilize the interactions among other components of the 20S particle. Cross-linking studies indicate that  $\alpha$ - and  $\gamma$ -SNAP interact directly when bound to membranes (S.W.W., unpublished results). Both proteins also interact directly with NSF, but only  $\alpha$ -SNAP seems to contact an integral membrane receptor protein<sup>9</sup>. The recombinant proteins expressed in *E. coli* were used as antigens to produce polyclonal anti-SNAP antibodies. Immunofluorescent localization of  $\alpha$ - and  $\gamma$ -SNAP using affinity-purified antibodies shows strong staining of intracellular membranes including, among others, the reticular endoplasmic reticulum (ER) and perinuclear Golgi (data not shown).

The role of NSF in intercisternal Golgi transport<sup>10</sup>, ER-Golgi transport<sup>11</sup>, endosome-endosome fusion<sup>12</sup> and transcytotic vesicle fusion with the plasma membrane<sup>13</sup> has clearly demonstrated its general function in the cell. The SNAPs mediating NSF binding to membranes appear to be equally widely distributed. Transport vesicles must therefore be directed to the appropriate target by some other mechanism, probably involving the SNAP receptor (SNARE) family described in the accompanying article<sup>4</sup>.

$\beta$ -SNAP may be the first example of a specialized SNAP. It is very similar in sequence to  $\alpha$ -SNAP, and competes for the same binding site on SNAP receptors<sup>8,9</sup>, but unlike  $\alpha$ -SNAP cannot restore cell-free Golgi transport activity to cytosol from *sec17* mutant yeast. Even in brain, it is less abundant than  $\alpha$ -SNAP. Its regionalized expression (with highest concentrations in areas of high neuron density<sup>6</sup>) suggests that it is specialized for certain types of neurosecretion. □

Received 25 February; accepted 9 March 1993.

- Weidman, P. J., Melançon, P., Block, M. R. & Rothman, J. E. *J. Cell Biol.* **108**, 1589–1596 (1989).
- Clary, D. O., Griff, I. C. & Rothman, J. E. *Cell* **61**, 709–721 (1990).
- Clary, D. O. & Rothman, J. E. *J. Biol. Chem.* **265**, 10109–10117 (1990).
- Sollner, T. et al. *Nature* **362**, 318–324 (1993).
- Griff, I. C., Schekman, R., Rothman, J. E. & Kaiser, C. A. *J. Biol. Chem.* **267**, 12106–12115 (1992).
- Kato, K. *Eur. J. Neurosci.* **2**, 704–711 (1990).
- Lupas, A., Van Dyke, M. & Stock, J. *Science* **252**, 1163–1164 (1991).
- Wilson, D. W., Whiteheart, S. W., Wiedmann, M., Brunner, M. & Rothman, J. E. *J. Cell Biol.* **117**, 531–538 (1992).
- Whiteheart, S. W., Brunner, M., Wilson, D. W., Wiedmann, M. & Rothman, J. E. *J. Biol. Chem.* **267**, 12239–12243 (1992).

- Block, M. R., Glick, B. S., Wilcox, C. A., Wieland, F. T. & Rothman, J. E. *Proc. natn. Acad. Sci. U.S.A.* **85**, 7852–7856 (1988).
- Beckers, C. J., Block, M. R., Glick, B. S., Rothman, J. E. & Balch, W. E. *Nature* **339**, 397–398 (1989).
- Diaz, R., Mayorga, L. S., Weidman, P. J., Rothman, J. E. & Stahl, P. D. *Nature* **339**, 398–400 (1989).
- Sztul, E., Colombo, M., Stahl, P. & Samanta, R. *J. Biol. Chem.* **268**, 1876–1885 (1993).
- Smith, T. F. & Waterman, M. S. *Appl. Math.* **2**, 488–489 (1981).

ACKNOWLEDGEMENTS. We thank A. Smith (Stanford University), J. Pohl (Emory University) for chemical sequencing of the SNAP peptides, and G. Tanigawa for advice. This work was supported by the Mathers Foundation and by an NIH grant (to J.E.R.), by a grant from The Jane Coffin Childs Memorial Fund for Medical Research (to S.W.W.), by an EMBO postdoctoral fellowship (to M.B.), and by a DFG fellowship (to T.M.).

## HIV infection is active and progressive in lymphoid tissue during the clinically latent stage of disease

Giuseppe Pantaleo\*, Cecilia Graziosi\*, James F. Demarest\*, Luca Butini†, Maria Montroni‡, Cecil H. Fox‡, Jan M. Orenstein§, Donald P. Kotler|| & Anthony S. Fauci\*

\* Laboratory of Immunoregulation, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20892, USA

† Department of Internal Medicine, University of Ancona, 60020 Ancona, Italy

‡ Department of Neuropathology, Yale University School of Medicine, New Haven, Connecticut 06510, USA

§ Department of Pathology, George Washington University, Washington DC 20037, USA

|| St Luke's-Roosevelt Hospital Center, New York, New York 10025, USA

**PRIMARY** infection with the human immunodeficiency virus (HIV) is generally followed by a burst of viraemia with or without clinical symptoms<sup>1–3</sup>. This in turn is followed by a prolonged period of clinical latency. During this period there is little, if any, detectable viraemia, the numbers of infected cells in the blood are very low, and it is extremely difficult to demonstrate virus expression in these cells<sup>4</sup>. We have analysed viral burden and levels of virus replication simultaneously in the blood and lymphoid organs of the same individuals at various stages of HIV disease. Here we report that in early-stage disease there is a dichotomy between the levels of viral burden and virus replication in peripheral blood versus lymphoid organs. HIV disease is active in the lymphoid tissue throughout the period of clinical latency, even at times when minimal viral activity is demonstrated in blood.

Viral burden was simultaneously analysed by DNA polymerase chain reaction (PCR) in mononuclear cells from

TABLE 1 Profiles of 12 patients in different stages of HIV infection who were analysed for viral burden

| Patient* | Sex | Age (years) | Transmission | Absolute CD4 count (cells per mm <sup>3</sup> ) | Tissue                         |
|----------|-----|-------------|--------------|-------------------------------------------------|--------------------------------|
| 1        | F   | 24          | Heterosexual | 571                                             | Axillary LN†                   |
| 2        | F   | 33          | Heterosexual | 620                                             | Cervical LN                    |
| 3        | M   | 32          | Homosexual   | 688                                             | Axillary LN                    |
| 4        | M   | 26          | IDU‡         | 707                                             | Axillary LN                    |
| 5        | M   | 28          | Homosexual   | 423                                             | Cervical LN                    |
| 6        | M   | 43          | Homosexual   | 355                                             | Cervical LN                    |
| 7        | M   | 30          | Homosexual   | 365                                             | Axillary LN                    |
| 8        | F   | 3           | Perinatal    | 1,798§                                          | Cervical LN, tonsils, adenoids |
| 9        | M   | 5           | Perinatal    | 913§                                            | Tonsils                        |
| 10       | M   | 43          | Homosexual   | 42                                              | Axillary LN                    |
| 11       | M   | 26          | IDU          | 194                                             | Axillary LN                    |
| 12       | F   | 37          | Heterosexual | 128                                             | Tonsils                        |

\* Lymph node biopsies were obtained under a protocol approved by the appropriate institutional review boards; no patients had an opportunistic infection at the time of the biopsy. Patients were divided into three groups on the basis of absolute count of CD4<sup>+</sup> T cells: early-stage disease (patients 1–4), more than 500 CD4<sup>+</sup> cells per mm<sup>3</sup>; intermediate-stage disease (patients 5–9) 200 to 500 CD4<sup>+</sup> T cells per mm<sup>3</sup>; late-stage disease (patients 10–12), less than 200 CD4<sup>+</sup> cells per mm<sup>3</sup>.

† LN, lymph node.

‡ IDU, intravenous drug user.

§ Patients 8 and 9 are paediatric patients; CD4<sup>+</sup> cell counts and clinical status are compatible with intermediate-stage disease.

peripheral blood and lymphoid tissues in 12 HIV-seropositive individuals who were divided into three groups according to the absolute count of CD4<sup>+</sup> T cells (Table 1). In agreement with our previous study<sup>5</sup>, in two representative patients in early and intermediate stages of disease, there is between 5 and 10 times greater frequency of infected cells in the lymphoid tissue (Fig. 1a, b). In patient 10 (advanced-stage disease), viral burden was equal in both peripheral blood and lymphoid tissue (Fig. 1c).

These results indicate a preferential localization of HIV-infected cells in the lymphoid tissue early in the course of HIV infection. This is associated with follicular hyperplasia that may be related to an abnormal movement of CD4<sup>+</sup> T lymphocytes, including those infected with HIV, resulting in sequestration of HIV-infected cells in the lymphoid organs.

The relative degree of virus replication was compared in lymphoid tissue mononuclear cells (LTMC) and peripheral blood mononuclear cells (PBMC) by RNA PCR. The sensitivity of the PCR assay and the linear relationship between the relative intensity of the PCR signals and different amounts of HIV RNA were determined first on RNA prepared from 10-fold serial cell

dilutions of 8E5, a chronically HIV-infected T-cell clone that constitutively expresses HIV<sup>6</sup> (Fig. 1g). Total RNA was prepared simultaneously from PBMC and LTMC, reverse-transcribed and amplified with primer pairs specific for *gag* and *tat/rev* gene segments. Striking differences in the levels of HIV RNA synthesis were consistently observed between PBMC and LTMC in all patients analysed, regardless of the stage of disease (Fig. 1d-f). In early and intermediate stages of disease (patients 1-9), expression of both structural and regulatory messages was barely detectable or undetectable in PBMC, whereas high levels of HIV-specific messages were detected in LTMC (Fig. 1d, e). In late-stage disease (patients 10-12), HIV RNA synthesis was

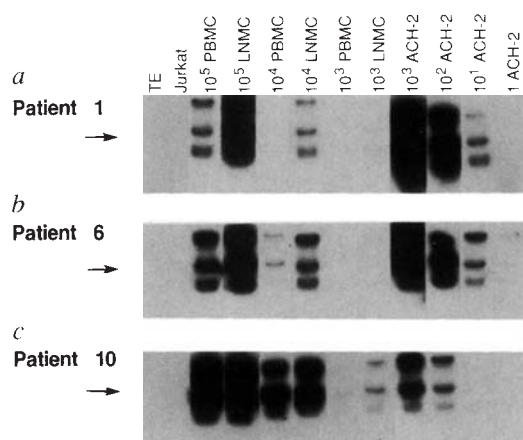
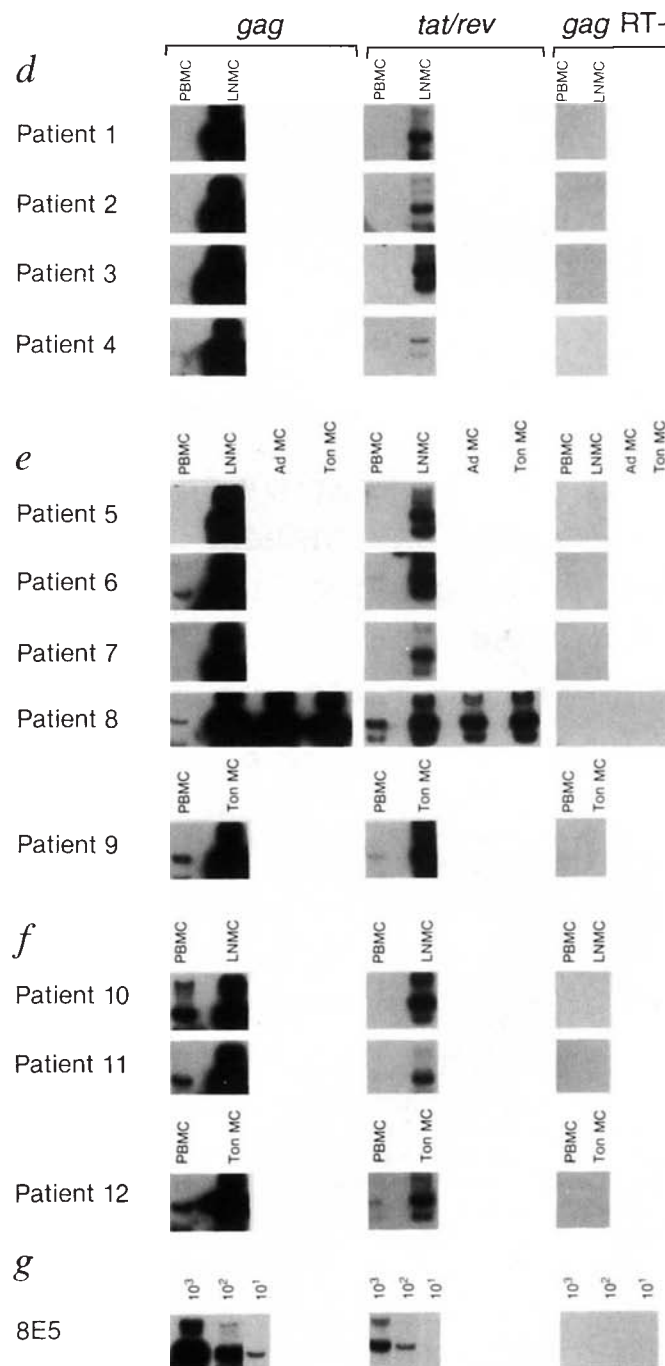


FIG. 1 *a-c*, PCR amplification of HIV-1 DNA in unfractionated cell populations isolated from peripheral blood (PB) and lymph node (LN) of three separate patients in different stages of disease. Intensities of the PCR signals in  $1 \times 10^5$  PBMC and LNMC are compared with 10-fold serial dilutions of cell lysates from ACH-2, a chronically infected T-cell clone containing one proviral copy per cell<sup>18</sup>. *a*, In patient 1, the frequency of HIV-infected cells is 1/10,000 in the PB, and between 1/100 and 1/1,000 in the LN. *b*, In patient 6, the frequency is between 1/1,000 and 1/10,000 in the PB, and 1/1,000 in the LN. *c*, In patient 10, the frequency is between 1/10 and 1/100 in both PB and LN. *d-g*, Analysis of HIV expression in PBMC and LTMC from patients in different stages of disease by RNA-PCR. *d*, Comparative analysis of expression of structural (*gag*) and regulatory (*tat/rev*) HIV-specific messages in PBMC and LNMC isolated from four patients (patients 1-4) in early-stage disease. *e*, Patients 5-9 were in intermediate-stage disease. In patients 5-7, HIV expression was comparatively analysed in PBMC and LNMC; in patient 8, similar analysis was done in MC isolated from PB, LN, adenoids (Ad) and tonsils (Ton) from the same patient, and in patient 9 in PBMC and TonMC. *f*, Patients 10-12 were in late-stage disease. A negative control (that is, RNA treated with DNase and amplified with primer pair specific for *gag* in the absence of reverse transcriptase) is shown for each patient. *g*, Analysis of expression of *gag* and *tat/rev* HIV-specific messenger RNA in 10-fold serial cell dilutions of 8E5.

**METHODS.** *a-c*, Isolation of MC from PB and LT and quantitative DNA PCR were done as previously described<sup>5</sup>. Briefly, for DNA PCR unfractionated cell populations isolated from PB and LN from the same patients were lysed, and dilutions were obtained by mixing different amounts of the stock lysate (corresponding to  $10^6$  cells) with Jurkat (leukaemia T-cell line) cell lysates (corresponding to  $10^5$  cells) in a total volume of 50  $\mu$ l. DNA corresponding to  $10^5$ ,  $10^4$  and  $10^3$  cells was amplified using a primer pair specific for the *gag* (SK145/101) gene segment, and amplified products were hybridized to the specific probe (SK102 for *gag*) that had been end-labelled using [ $\gamma$ -<sup>32</sup>P]ATP. Products of hybridization were analysed by electrophoresis in 10% polyacrylamide gels and visualized by autoradiography. As a positive control, 10-fold serial dilutions of ACH-2 were used; as a negative control lysis buffer (LB) and cell lysates from  $10^5$  Jurkat cells were used. *d-g*, RNA-PCR was done as previously described<sup>19</sup>. To determine the levels of HIV RNA synthesis in PB and LT, PBMC and LTMC were isolated simultaneously from the PB and LT from the same patients, and pelleted in aliquots of  $3 \times 10^6$  cells per vial and stored at  $-70^\circ\text{C}$  until used. Total RNA was extracted by the RNeasy (Qiagen/Biotech Laboratories, Friendswood, Texas) method<sup>20</sup>. RNA was treated with an excess of RNase-free DNase. Two aliquots of each sample were reverse-transcribed and amplified with primer pairs specific for *gag* (SK 38/39)<sup>21,22</sup> and *tat/rev* (TR-5/TR-3)<sup>23</sup>. A third aliquot was



amplified with a primer pair specific for *gag*, but not reverse-transcribed (negative control). Amplified products were hybridized to the specific probes (SK 19 for *gag* and TR-4 for *tat/rev*) that had been end-labelled using [ $\gamma$ -<sup>32</sup>P]ATP, analysed by electrophoresis in 10% polyacrylamide gels and visualized by autoradiography.



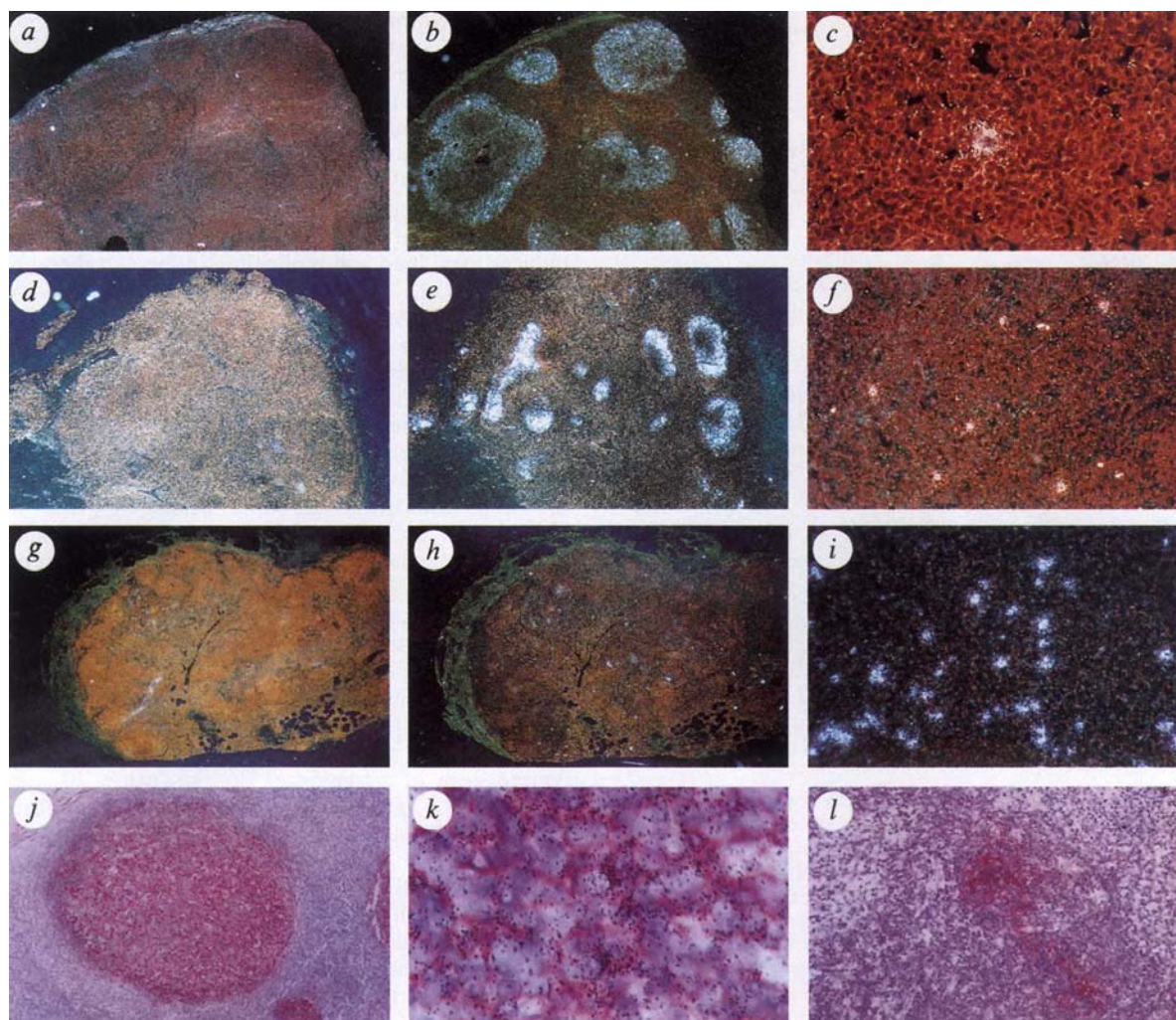


FIG. 2 *In situ* hybridization of lymph node sections from HIV-infected patients in different stages of disease. *a-c*, Early-stage disease. *a*, Darkfield image of lymph node section not digested with protease. *b*, After protease digestion. Location of HIV RNA is indicated by the silver grains, which appear as white dots. An intense, granular, diffuse signal is predominantly restricted to the area of the germinal centres. *c*, Darkfield image without protease digestion. Localization of HIV RNA signal over an individual cell in the paracortical area. *d-f*, Patient with intermediate-stage disease. *d*, Darkfield image without protease digestion. *e*, After protease digestion. Intense and diffuse signal is present over some germinal centres, whereas in the same section the signal is significantly decreased in other germinal centres. *f*, HIV RNA signal over individual cells in the paracortical area. *g-i*, Patient with late-stage disease. *g*, Darkfield image without protease digestion. Lymph node architecture is completely disrupted, and germinal centres are virtually absent. *h*, After protease digestion. A moderate increase in *in situ* hybridization signal is seen. *i*, Higher magnification of a protease-digested section showing distribution of silver grains over several isolated cells expressing HIV RNA in copious amounts. *j*, Brightfield image of a protease-digested lymph node section subjected to immunohistochemistry plus *in situ* hybridization showing the staining (New Fuchsin Red) of the FDC network with

anti-CD21 antibody. *k*, Localization of HIV RNA (that is, with silver grains which appear as black dots) coincides with the staining of the FDC network. *l*, Presumed already involuted germinal centre with a remnant of FDC network. Data are representative of the qualitative changes as well as the degrees of changes among the different stages of disease seen in the 7 of 12 patients in the present study in whose lymph nodes *in situ* hybridization was done. These data were consistent with data obtained from an additional 30 patients in whose lymph nodes similar *in situ* hybridization studies were done.

**METHODS.** *In situ* hybridization was performed as previously described<sup>8</sup>. RNA probes were synthesized from five DNA templates that represent 90% of the HIV-1 genome<sup>8</sup>. Tissues were fixed in formaldehyde, embedded in paraffin and digested with proteolytic enzymes. Slides were hybridized with sense or antisense probes and run in duplicate. After hybridization, autoradiograms were made with NTB3 emulsion (Eastman Kodak), stained with haematoxylin-eosin, and examined by darkfield microscopy. In double-labelling (immunohistochemistry plus *in situ* hybridization) experiments, slides were subjected first to staining with anti-CD21 antibody. Anti-CD21 antibody (clone 1F8) that specifically stains FDC in formaldehyde-fixed tissues was purchased from Dako Corporation, Carpinteria, CA.

increased in PBMC compared to that in earlier stages of disease (Fig. 1*f*), although it still remained clearly lower than in LTMC. Similar patterns of HIV expression in PBMC of patients at different stages of disease have been previously observed<sup>7</sup>. But this is the first demonstration of a striking dichotomy between peripheral blood and lymphoid tissue from the same individuals in viral burden and replication. Of note, high levels of HIV expression were observed in lymphoid tissue other than lymph nodes, such as adenoids and tonsils (Fig. 1*e, f*), indicating a systemic dissemination of HIV among lymphoid tissue and not

a localization to lymph nodes.

In order to delineate the pathogenic mechanisms responsible for the dichotomy between the levels of viral burden and expression in peripheral blood compared to lymphoid tissue, we analysed the latter by *in situ* hybridization and transmission electron microscopy. All lymph nodes obtained from patients in early-stage disease had some degree of follicular hyperplasia. In these lymph nodes, after protease treatment, most of the hybridization signal, which corresponded to extracellular viral particles<sup>8</sup>, was restricted to the germinal centres (Fig. 2*b*). The



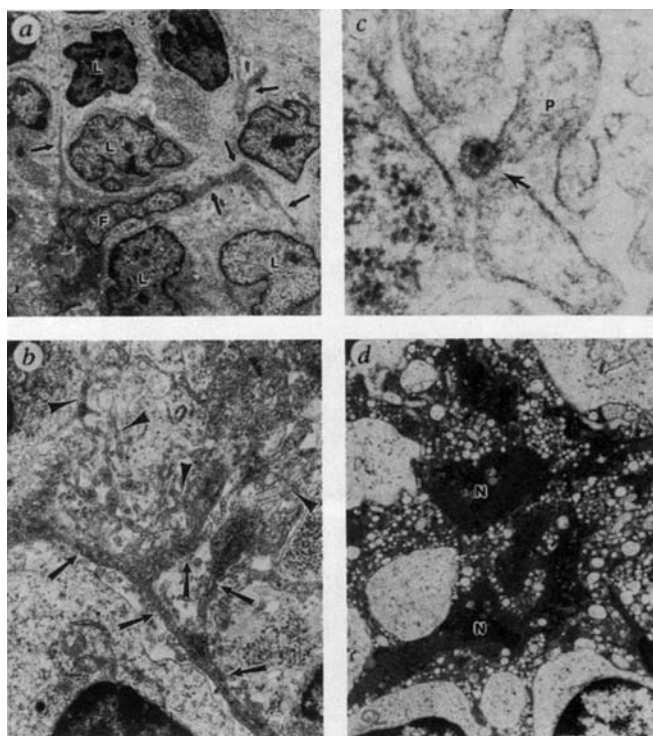


FIG. 3 *a*, Electron-dense FDC (for example, F) extends processes (for example, arrows) that contact several adjacent lymphocytes (for example, L). Magnification,  $\times 1,920$ . *b*, Moderately dense extensions (for example, arrows) of the FDC branch several times before ending in numerous villus processes (for example, arrowheads). Magnification,  $\times 6,720$ . *c*, A mature HIV particle (for example, arrow), with its typical conical nucleoid, located in the intercellular space of a hyperplastic germinal centre. It is closely associated with FDC processes (for example, P) and is partially obscured by filamentous material. Magnification,  $\times 55,200$ . *d*, Several tightly clustered FDC have irregular, markedly condensed nuclei (for example, N) and electron-dense, vesiculated cytoplasm which are signs of degeneration and necrosis. Magnification,  $\times 3,744$ . Micrographs were obtained from LN specimens of the patients listed in Table 1. Data are representative of the findings in the 5 of 12 patients in this study on whose lymph nodes electron microscopy was done. These data are consistent with an additional eight patients whose lymph nodes were similarly examined by electron microscopy.

paucity of the signal in the absence of protease treatment (Fig. 2*a*) was consistent with extracellular virions coated with proteins, forming immune complexes<sup>8</sup>. According to previous studies<sup>8-15</sup>, and as confirmed by ultrastructural analysis, HIV particles are trapped on the villus processes of follicular dendritic cells (FDC) (Fig. 3*c*) that surround and are intimately associated with lymphocytes (Fig. 3*a, b*). As disease progresses, important changes in the distribution pattern of HIV in the lymph nodes occur. During intermediate-stage disease, within a given lymph node, virus is trapped in some germinal centres and not in others (Fig. 2*e*). In late-stage disease, the architecture of the lymph node is disrupted and most of the germinal centres are involuted concomitant with a loss of virus-trapping capability of the node (Fig. 2*g, h*). Degeneration and death of FDC, as indicated by ultrastructural analysis (Fig. 3*d*), is generally associated with disease progression. But varying degrees of histopathology can be seen focally in the lymph nodes of patients even in early stages of disease. Furthermore, isolated cells expressing HIV RNA are detected in the paracortical area regardless of the stage of disease, and the proportion of these cells seems to increase with disease progression (Fig. 2*c, f, i*).

The mechanism responsible for the trapping of virus was directly addressed by combining *in situ* hybridization for HIV with immunohistochemistry for FDC using anti-CD21 antibody. Figure 2*j* illustrates an intact dense FDC network in the germinal

centre of a lymph node from a patient in early-stage disease. It is notable that the hybridization signal for HIV RNA, which corresponds to extracellular HIV particles, completely overlaps with the FDC network (Fig. 2*k*). In lymph nodes of patients in late-stage disease, only isolated FDC are detected (Fig. 2*l*), and virus trapping is virtually absent although the residual FDC can still retain virus (data not shown). These results indicate that efficient trapping of virus corresponds directly with the integrity of the FDC network. Thus, the mechanisms responsible for reduced viral load and replication in the peripheral blood compartment in the early stages of disease involve a combination of the mechanical filtering and trapping of virions by the FDC network and the sequestration of infected CD4<sup>+</sup> T cells in the hyperplastic lymph node. Furthermore, the formation of immune complexes (virus plus immunoglobulin or complement) would contribute to the attachment of HIV to the FDC. In this regard, it is likely that an initially adequate immune response itself also contributes substantially to the clearance of virus from the circulation. In the late stages of disease<sup>4</sup>, these mechanical mechanisms are altered and an effective immune response against HIV is lost<sup>16</sup>, leading to an increase in viral burden in PBMC and an acceleration of plasma viraemia, explained at least in part by the spillover of virus from the lymphoid organs.

The observations of low to absent levels of viraemia and virus replication in PBMC have led to the impression that HIV disease is not active during the period of clinical latency. This has contributed to the general policy of initiating antiretroviral therapy only at the time of the appearance of constitutional symptoms or of a significant decline (below 500 per mm<sup>3</sup>) of CD4<sup>+</sup> T cells<sup>17</sup>. Here we have demonstrated that during clinical latency, HIV accumulates in the lymphoid organs and replicates actively despite a low viral burden and low-to-absent viral replication in PBMC. Therefore, a state of true microbiological latency does not exist during the course of HIV infection. The peripheral blood does not accurately reflect the actual state of HIV disease, particularly early in the clinical course of HIV infection. In fact, HIV disease is active and progressive even when there is little evidence of disease activity by readily measured viral parameters in the peripheral blood, and the patient is experiencing clinical latency. These findings not only advance our understanding of the immunopathogenic mechanisms of HIV infection, but may have important implications in the design of therapeutic strategies. □

Received 18 November 1992; accepted 25 February 1993.

- Daar, E. S., Moudgil, T., Meyer, R. D. & Ho, D. D. *New Engl. J. Med.* **324**, 961-964 (1991).
- Clark, S. J. *et al. New Engl. J. Med.* **324**, 954-960 (1991).
- Tindall, B. & Cooper, D. A. *AIDS* **5**, 1-14 (1991).
- Fauci, A. S., Schnittman, S. M., Poli, G., Koenig, S. & Pantaleo, G. *Ann. Intern. Med.* **114**, 678-693 (1991).
- Pantaleo, G. *et al. Proc. natn. Acad. Sci. U.S.A.* **88**, 9838-9842 (1991).
- Folks, T. M. *et al. J. exp. Med.* **164**, 280-290 (1986).
- Michael, N. L., Vahey, M., Burke, D. S. & Redfield, R. R. *J. Virol.* **66**, 310-316 (1992).
- Fox, C. H. *et al. J. infect. Dis.* **164**, 1051-1057 (1991).
- Racz, P. *et al. Prog. Allergy* **37**, 81-181 (1986).
- Armstrong, J. A., Dawkins, R. L. & Horne, R. *Immun. Today* **6**, 121-122 (1985).
- Biberfeld, P. *et al. Acta path. microbiol. immun. scand.* **95**, 47-65 (1987).
- Armstrong, J. A. in *Accessory Cells in HIV and Other Retroviral Infections* (eds Racz, P., Dijkstra, C. D. & Gluckman, J. C.) 69-82 (Karger, Basel, 1991).
- Spiegel, H., Herbst, H., Niedobitek, G., Foss, H. D. & Stein, H. *Am. J. Path.* **140**, 15-22 (1992).
- Emilie, D. *et al. J. clin. Invest.* **86**, 148-159 (1990).
- Tenner-Racz, K., Racz, P., Dietrich, M. & Karin, P. *Lancet* **i**, 105-106 (1985).
- Haynes, B. F. in *AIDS Etiology, Diagnosis, Treatment and Prevention* (eds DeVita, V. T. Jr *et al.*) 77-86 (Lippincott, Philadelphia, 1992).
- Volberding, P. A. *et al. New Engl. J. Med.* **322**, 941-949 (1990).
- Clouse, K. A. *et al. J. Immun.* **142**, 431-438 (1989).
- Graziosi, C., Pantaleo, G. & Fauci, A. S. *Curr. Protoc. Immun.* (in the press).
- Chomczynski, P. & Sacchi, N. *Analyt. Biochem.* **162**, 156-159 (1987).
- Kwok, S. *et al. J. Virol.* **61**, 1690-1694 (1987).
- Ou, C. Y. *et al. Science* **239**, 295-297 (1988).
- Schnittman, S. M., Greenhouse, J. J., Lane, H. C., Pierce, P. F. & Fauci, A. S. *AIDS Res. Hum. Retrovir.* **7**, 361-367 (1991).

ACKNOWLEDGEMENTS. We thank P. A. Pizzo for providing the paediatric patients, M. A. Rust for editorial assistance, and S. Rosenberg and staff of the National Cancer Institute Surgery Branch for cooperation. This work counts in partial fulfillment of the Ph.D. requirements for J.F.D. in the Graduate Genetics Program at the George Washington University. Partly supported by AIDS Project, Istituto Superiore Sanita', Italy.



# Massive covert infection of helper T lymphocytes and macrophages by HIV during the incubation period of AIDS

Janet Embretson\*, Mary Zupancic\*, Jorge L. Ribas<sup>†‡</sup>, Allen Burke<sup>‡</sup>, Paul Racz<sup>§</sup>, Klara Tenner-Racz<sup>§</sup> & Ashley T. Haase\*||

\* Department of Microbiology, University of Minnesota,

420 Delaware Street S.E., Minneapolis, Minnesota 55455, USA

<sup>†</sup> Henry M. Jackson Foundation, Pockville, Maryland, 20850, USA

<sup>‡</sup> Department Cardiovascular Pathology, Armed Forces Institute of

Pathology, Department of Defense, Washington DC 20306, USA

<sup>§</sup> Bernhard-Nocht Institute for Tropical Medicine,

Bernhard-Nocht Strasse 74, 2000 Hamburg 36, Germany

|| To whom correspondence should be addressed

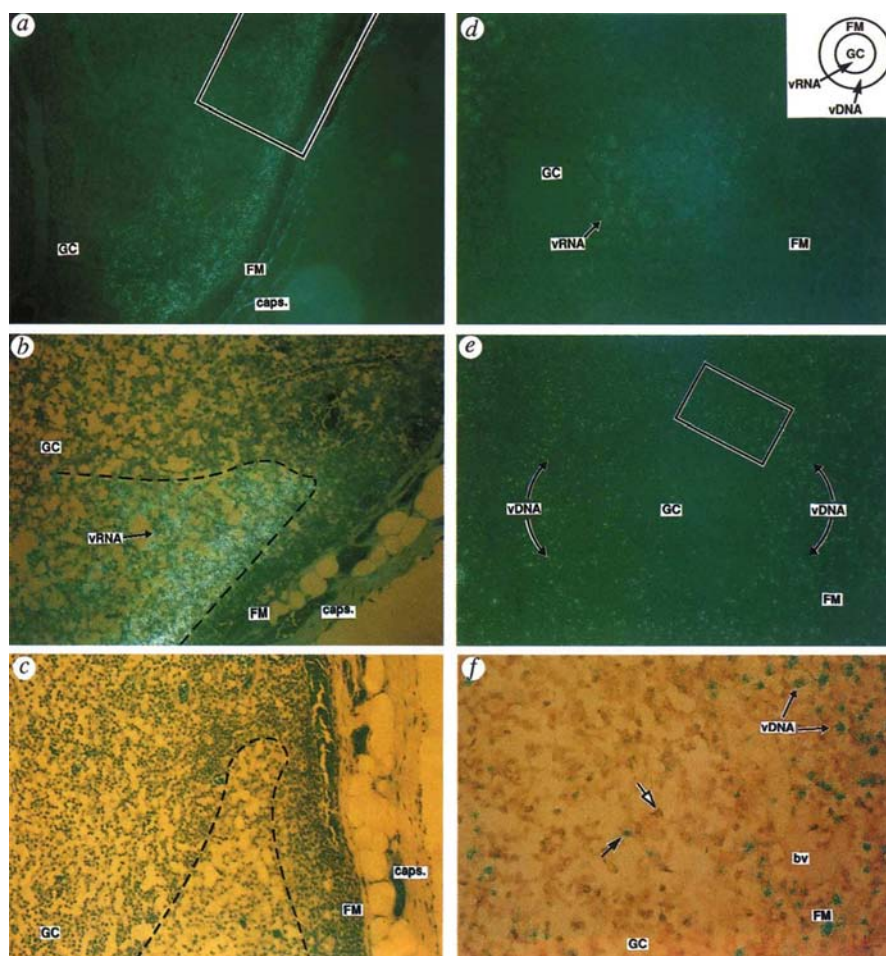
**ANIMAL and human lentiviruses elude host defences by establishing covert infections and eventually cause disease through cumulative losses of cells that die with activation of viral gene expression<sup>1-5</sup>. We used polymerase chain reaction *in situ* double-label methods<sup>6,7</sup> to determine how many CD4<sup>+</sup> lymphocytes are latently infected by human immunodeficiency virus (HIV) in patient lymph nodes and whether the pool of infected cells is large enough to account for immune depletion through continual activation of viral gene expression and attrition of cells responding to**

antigens. We discovered an extraordinarily large number of latently infected CD4<sup>+</sup> lymphocytes and macrophages throughout the lymphoid system from early to late stages of infection, and confirmed<sup>8-14</sup> the extracellular association of HIV with follicular dendritic cells. Follicular dendritic cells may transmit infection to cells as they migrate through lymphoid follicles. Latently infected lymphocytes and macrophages constitute an intracellular reservoir large enough ultimately to contribute to much of the immune depletion in AIDS, and represent a difficult problem that must be resolved in developing effective treatments and protective vaccine.

We applied polymerase chain reaction (PCR) *in situ* and *in situ* hybridization techniques to define the anatomic distribution, number and types of cells containing HIV DNA and RNA in lymphoid tissues of four individuals (Table 1a). Cases 1-3 were presumably asymptomatic with no histopathological evidence of opportunistic infections or neoplasms. Case 4 represents a late stage of infection (CDC IV.D).

The distribution of cells with HIV DNA and/or RNA fell into two patterns. In case 2, viral RNA was detected in the germinal centres in virions overlying the processes of follicular dendritic cells (FDCs), inferred from the diffuse distribution of the hybridization signal shown previously<sup>13</sup> to be characteristic of this kind of association of HIV with FDCs. There was no detectable viral RNA or DNA in individual cells in germinal centres, follicular mantle or paracortex (Fig. 1a-c; Table 1b). In the other cases, HIV RNA was also primarily in germinal centres, mainly in the apical light zone, but we did detect HIV DNA in cells predominantly around the germinal centres (Fig. 1d-f). We did not detect in any case viral DNA in FDCs with PCR *in situ* techniques sensitive enough to identify cells with

FIG. 1 Distribution of HIV RNA and DNA in lymph nodes. **a**, Viral RNA is in a germinal centre (GC) adjacent to the follicular mantle zone (FM) and capsule (caps); magnification,  $\times 18$ . With epipolarized illumination and filters, silver grains appear white or green. **b**, Higher magnification ( $\times 36$ ) of insert in **a**. Hybridization signal is diffuse, characteristic of vRNA associated with the cytoplasmic processes of FDCs. **c**, None of the cells has vDNA, as no hybridization signal is detectable in the area delineated by the dashed line. **d-f**, In other cases, CD4<sup>+</sup> lymphocytes contain HIV DNA. **d**, **e**, vRNA is associated with FDCs in the GC and vDNA is in cells in the FM, as diagrammed in the insert in **d** (**d**, **e**, magnification  $\times 21$ ). **f**, Area of insert in **e** enlarged ( $\times 53$ ). Combined epipolarized and transillumination provides contrasting signal for HIV DNA (green-silver grains) over nuclei with the brown-stained cytoplasm of CD4<sup>+</sup> lymphocytes. There are some CD4<sup>+</sup> cells with HIV DNA in the FM, surrounding a blood vessel (bv), and in the GC. Infected cells (small arrow) adjoin uninfected CD4<sup>+</sup> cells with no signal over the nucleus (large arrow). **METHODS.** vRNA was detected by *in situ* hybridization<sup>21</sup> with a radiolabelled HIV-specific probe. vDNA was detected by PCR amplification and hybridization *in situ*<sup>4-8</sup>. Sections were amplified for 40 cycles with primers from the *gag* region of HIV. After amplification, sections were hybridized to HIV-specific (BH10 DNA) or a heterologous visna virus probe radiolabelled to comparable specific activity. Radioautographs were developed after 4-11 days of exposure. For double-label PCR *in situ*, sections were treated to block endogenous peroxidase activity, briefly digested with pronase E and incubated with monoclonal antibody for CD4<sup>+</sup> cells (OPD4, Dako). Following secondary antibody, avidin-biotin peroxidase complex and diaminobenzidine H<sub>2</sub>O<sub>2</sub> reactions to deposit chromogen on the CD4<sup>+</sup> lymphocytes<sup>5,7</sup>, sections were carried through the PCR *in situ* procedures already described, hybridized, washed, coated with NTB-2 and developed and stained after 16 days.





a single copy of proviral DNA, nor did we see viral particles budding from FDCs. Although there were many viral particles on the surfaces of FDCs (Fig. 2) in the interdendritic extracellular spaces<sup>4-6,15</sup>, we think that with rare exceptions<sup>8</sup> FDCs are not replicating HIV, but trap virions on their membranes in

immune complexes (electron-dense material in figure) or through complement receptors.

We showed by a double-label PCR *in situ* method that most cells with HIV DNA in lymph nodes are CD4<sup>+</sup> T helper lymphocytes (Figs 1f and 3a). The radioautographic signal was local-

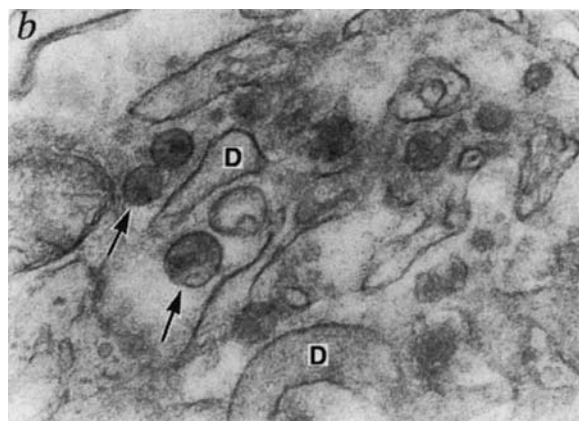
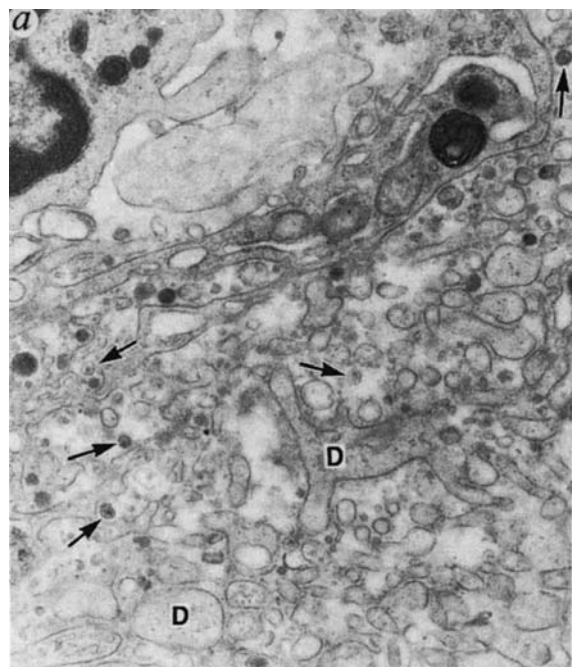


FIG. 2 HIV particles in the interdendritic spaces. *a*, Extracellular viral particles (arrows) embedded in moderately electron-dense material shown previously<sup>10</sup> to be virus immune complexes covering dendritic processes (*d*) of FDCs ( $\times 19,000$ ). *b*, Viral particles in more detail ( $\times 30,000$ ).

FIG. 3 HIV DNA and RNA in CD4<sup>+</sup> lymphocytes and macrophages. *a*, HIV DNA in CD4 lymphocytes identified by staining with antibody to CD4 followed by PCR *in situ*. The cytoplasm of CD4<sup>+</sup> lymphocytes is stained brown. Cells with vDNA have green-appearing silver grains over their nuclei. Because the grains are in a higher focal plane the cells are slightly out of focus. In *a*, there are three representative panels of high power fields (magnification  $\times 165$ ) from two cases; small arrows, a CD4<sup>+</sup> lymphocyte with HIV DNA; large arrows, cells with no vDNA. There also are cells with vDNA (small arrows) adjacent to negative T4 cells (large arrows) or unidentified stromal elements (○) in *a*. *b*, No HIV DNA was detectable with the double-label procedure in seronegative control tissues (magnification  $\times 207$ ); or *c*, in many areas of nodes from seropositive cases after PCR *in situ* alone. *d*, Lymphocytes and monocytes/macrophages with vRNA at high copy number (transillumination only, silver grains appear black;  $\times 165$ . Small arrow, cells with abundant vRNA; case 4). *e*, None of the macrophages (MP) lining venous sinuses (VS) in the red pulp (rp) in this section of spleen had detectable vRNA (magnification  $\times 33$ ). *f*, At higher magnification ( $\times 83$ , in the region delineated by—in *e*) in a subjacent section, many MP lining the sinuses contain vDNA (cells with green grains over their nuclei). The cells were identified by morphology, characteristic location and staining of adjacent section with specific antibodies (anti-CD68). Methods as described in the legend to Fig. 1.

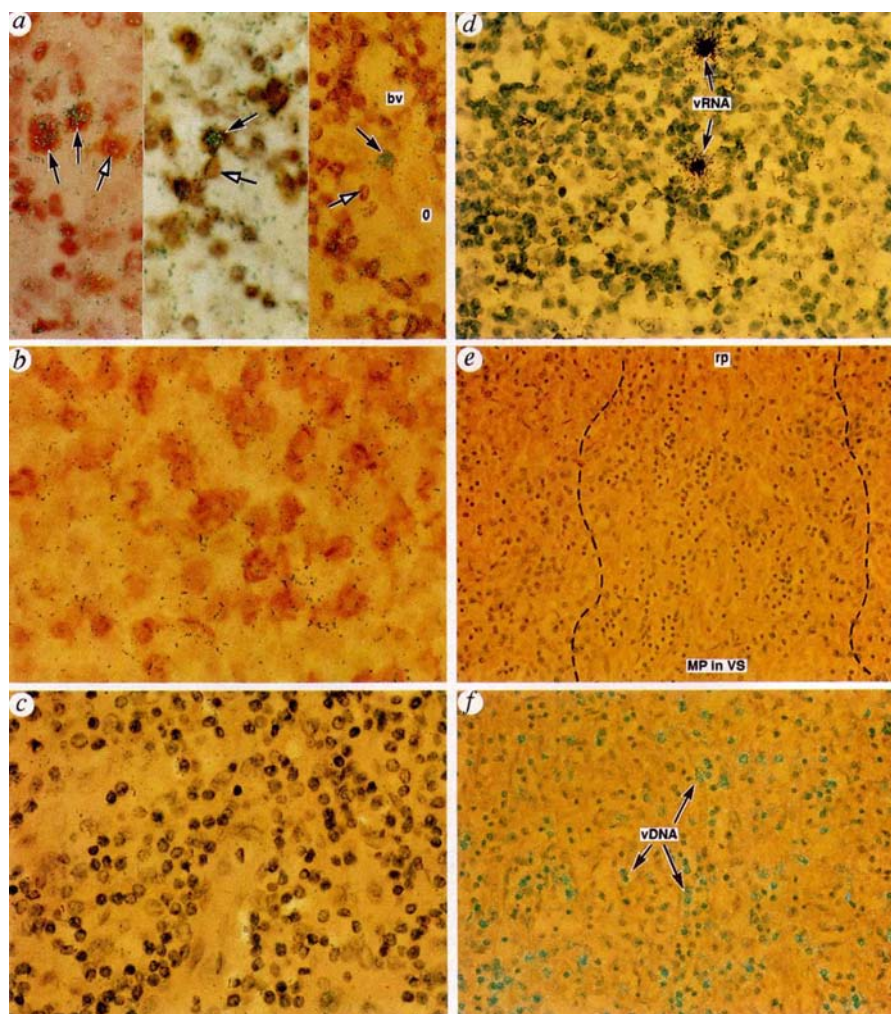




TABLE 1 Case descriptions and frequency of cells with viral DNA and RNA

| (a) Case description of HIV-1-infected individuals whose lymphoid tissues were evaluated |                |              |                                                                                                                                                       |                                  |
|------------------------------------------------------------------------------------------|----------------|--------------|-------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------|
| Case number                                                                              | Age (yrs), sex | Transmission | Clinicopathological status                                                                                                                            | Therapy                          |
| 1                                                                                        | 40, M          | IV drug use  | FH, some normal GC, no OI                                                                                                                             | None                             |
| 2                                                                                        | 28, M          | IV drug use  | FH, no evidence of OI                                                                                                                                 | None                             |
| 3                                                                                        | 37, F          | IV drug use  | FH, no evidence of OI                                                                                                                                 | None                             |
| 4                                                                                        | 32, M          | Homosexual   | Candidiasis; Kaposi sarcoma; CD4/CD8 in peripheral blood 105/480=0.2; LN: FH, hypervascular, increased plasma cells in the extrafollicular parenchyma | AZT (3'-azido-3'-deoxythymidine) |

(b) Extent of infection and viral gene expression in lymph nodes

| Case | Lymph node                                             | Fraction of CD4 <sup>+</sup> lymphocytes with HIV DNA |                      | Fraction of lymphocytes and macrophages with viral RNA |                      |
|------|--------------------------------------------------------|-------------------------------------------------------|----------------------|--------------------------------------------------------|----------------------|
|      |                                                        | Germinal centres and follicular mantle                | Paracortical regions | 30–60 copies                                           | 250–500 copies       |
| 1    | Axillary (mediastinal, supraclavicular)                | 0.32 (0.17–0.56)                                      | 0.21 (0.22–0.37)     | 0.06                                                   | 6 × 10 <sup>−4</sup> |
| 2    | Axillary (mesenteric, inguinal)                        | 0                                                     | 0                    | <10 <sup>−5</sup>                                      | <10 <sup>−5</sup>    |
| 3    | Axillary (mediastinal, supraclavicular, Peyer's patch) | 0.31 (0.29–0.34)                                      | 0.34 (0.32–0.37)     | 0.05                                                   | 7 × 10 <sup>−4</sup> |
| 4    | Axillary                                               | 0.21 (0.12–0.30)                                      | 0.16 (0.10–0.27)     | 0.03                                                   | 8 × 10 <sup>−4</sup> |

Lymph nodes and spleen were obtained at the time of autopsy from three presumably asymptomatic individuals who died of drug overdose and who were found to be seropositive for HIV-1 at that time. The pathological status was determined from post-mortem examination. A lymph node biopsy was obtained for diagnostic purposes from case 4, a seropositive patient in the late symptomatic stage of infection undergoing treatment with AZT. Lymph nodes were fixed in 4% paraformaldehyde or 10% buffered formalin 11, 16 and 17 hours after death in cases 1-3, and immediately after biopsy in case 4. The frequency of T lymphocytes with HIV DNA was rigorously determined with the double-label PCR *in situ* method described in the legend to Fig. 3 in and around three GCs selected randomly in the tissue section. Two fields in the GCs, FM and PC were again randomly selected at low magnification and the frequency of HIV-positive CD4 cells (cells with brown cytoplasm and  $\geq 10 \times$  number of silver grains in their nuclei compared to the seronegative control) per 100 cells was determined. The average and range are shown for each case and lymph node in parentheses. Results were qualitatively equivalent (localization and frequency of cells with viral DNA) in the node listed and those shown in parentheses. These were examined in two independent experiments preceding the development of the double-label PCR *in situ* technique. The frequency of lymphocytes and macrophages with 250-500 copies of viral RNA was assessed in the same GCs and paracortical regions in subjacent sections by *in situ* hybridization, scoring cells like the ones shown in Fig. 3d as positive. But because of the low frequency in these cells, 25-70 high-power microscopic fields (magnification,  $\times 500$ ) with 10,000-15,000 cells had to be examined to find 5-10 positive cells. In case 2, the entire lymph node was surveyed without detecting vDNA or vRNA in individual cells, although there was abundant vRNA associated with FDCs in most GCs in the sections listed (see Fig. 1a-b). The frequency of cells with 30-60 copies per cell was determined by comparing the frequency of grain counts for 100 cells in subjacent sections to those in which the frequency of cells with vDNA had been determined to grain counts in 100 cells in sections from lymphoid tissue (node, tonsil) of two seronegative controls hybridized *in situ* to detect vRNA. Cells in the infected lymph node with grain counts that fell outside the control distribution were classified as infected cells with HIV RNA ( $P \geq 0.999$ ). From the relationship between grain counts and the number of copies of vRNA per cell described in ref. 5, the number of copies of HIV RNA per cell was estimated to be 30-60 copies per cell. FH, Follicular hyperplasia; OI, opportunistic infections; GC, germinal centre; FM, follicular mantle; PC, paracortex.

ized to individual nuclei of immunocytochemically marked CD4<sup>+</sup> cells, with no evidence of leakage of PCR product to adjacent cells where it might be amplified to give falsely high estimates of the frequency of cells containing HIV DNA. Additional controls that collectively substantiate the specificity of the amplification and detection methods, and the accuracy of the estimates of the viral burden (number of infected cells) included: (1) no signal over nuclei in lymph nodes of seronegative individuals after amplification and hybridization with an HIV-specific probe (Fig. 3b); (2) an internal control, shown in Fig. 3c, of an area lacking cells with HIV DNA in the same section, shown in Fig. 1d, e, where there were many cells with HIV DNA; (3) in seropositive individuals, no signal over background in sections hybridized to HIV-specific probes without amplification; or (4) binding of heterologous probe (visna virus) substituted in the *in situ* hybridization step after amplification with HIV-specific primers (not shown).

The fraction of lymphocytes with viral DNA varied but was consistent in the same lymph node in three independent experiments. The range of estimated frequencies (Table 1b) of CD4<sup>+</sup> lymphocytes that contain HIV DNA, unequivocally determined with the double-label PCR *in situ* technique, is relatively narrow and consistent within and between cases. The averages indicate high levels of infection of CD4<sup>+</sup> lymphocytes in and around many germinal centres, in paracortical areas, and around blood vessels (Fig. 1f).

Most CD4<sup>+</sup> lymphocytes must be latently infected because we did not detect viral RNA in more than a small fraction of the cells in which HIV DNA was detectable. One in 100-400 cells with HIV DNA had viral RNA (Fig. 3d; Table 1b) as

abundant as productively infected cells<sup>5,6</sup>; and a higher proportion of cells had low levels of viral RNA (Table 1b). Similarly, HIV frequently establishes latent infections in monocytes and macrophages. In lymph node, we occasionally detected HIV DNA in monocytes/macrophages, but in the spleen of case 1 there were large numbers of infected macrophages in the venous sinuses of the red pulp (Fig. 3e, f). In the white pulp, cells with HIV DNA followed the distribution of T lymphocytes around the central arteriole and follicular mantle and germinal centre (not shown). Again, HIV RNA was generally not detectable in macrophages or lymphocytes (Fig. 3), in accord with previous reports<sup>16</sup>.

This analysis of the number and type of cell that HIV infects in lymphoid tissues, and discovery of widespread covert infections in macrophages and CD4<sup>+</sup> lymphocytes at early stages brings a fresh perspective to understanding the pathogenesis of disease and an increased appreciation of difficulties in treating and preventing infection. We think that very early in infection HIV is disseminated throughout the lymphoid system and sequestered in macrophages or trapped by FDCs, first through complement receptors and subsequently in immune complexes<sup>8-14</sup>. The macrophage tropism of early viral isolates<sup>17,18</sup> may reflect transmission of infection by cells of this lineage or be a consequence of a bottleneck for the quasi-species in which HIV first encounters and must infect macrophages. The one case in which the lymph node had HIV RNA associated with FDCs but no cell with HIV DNA, we interpret as evidence that HIV is sequestered first by FDCs and that at later stages virus is transmitted to CD4<sup>+</sup> lymphocytes as they migrate through the germinal centre to the follicular mantle and paracortical areas.

We have now shown that there are large numbers of latently infected CD4<sup>+</sup> lymphocytes and macrophages in which HIV can persist and be disseminated despite immune surveillance. The extent of infection in the CD4<sup>+</sup> lymphocyte population in lymphoid organs where 98% of these cells reside is sufficient to account for a substantial portion of the immune depletion in AIDS by a mechanism of slow elimination of small fractions of cells in which productive infection is continually activated concomitantly with the immune response to a particular antigen or pathogen<sup>19</sup>. That the fate of so many CD4<sup>+</sup> lymphocytes is determined relatively early in the course of infection underscores the difficulties in designing treatment programmes that can be instituted in time to influence the outcome favourably. Equally difficult is the problem of designing a vaccine to interdict the spread of HIV by a 'Trojan horse' mechanism<sup>20</sup>, concealed inside covertly infected cells. It is perhaps more realistic to hope that these interventions will limit infection to the extent necessary to prevent or moderate disease. □

Received 31 December 1992; accepted 2 March 1993.

1. Haase, A. T. *Nature* **322**, 130–136 (1986).
2. Fauci, A. S. *Science* **239**, 617–622 (1988).
3. Brähic, M., Stowring, L., Ventura, P. & Haase, A. T. *Nature* **292**, 240–242 (1981).
4. Staskus, K. A. et al. *Microbial Pathogen* **11**, 67–76 (1991).
5. Embretson, J. E. et al. *Proc. natn. Acad. Sci. U.S.A.* **90**, 357–361 (1993).
6. Haase, A. T., Retzel, E. F. & Staskus, K. A. *Proc. natn. Acad. Sci. U.S.A.* **87**, 4971–4975 (1990).
7. Brähic, M. & Haase, A. T. *Curr. Top. microbiol. Immun.* **143**, 9–20 (1989).
8. Armstrong, J. A. & Horne, R. *Lancet* **ii**, 370–372 (1984).
9. Racz, P. et al. *Progr. Allergy* **37**, 81–181 (1986).
10. Tenner-Racz, K. et al. *Am. J. Path.* **123**, 9–15 (1986).
11. Biberfeld, P. et al. *Am. J. Path.* **123**, 436–442 (1986).
12. Baroni, C. D., Pezzella, F., Miolo, M., Ruco, L. P. & Rossi, G. B. *Histopathology* **10**, 5–13 (1986).
13. Fox, C. H. et al. *J. infect. Dis.* **164**, 1051–1057 (1991).
14. Spiegel, H., Herbst, H., Niedobitek, G., Foss, H.-D. & Stein, H. *Am. J. Path.* **140**, 15–22 (1992).
15. Nuovo, G. J., MacConnell, P., Forde, A. & Delvenne, P. *Am. J. Path.* **139**, 847–854 (1991).
16. Harper, M. E., Marselle, L. M., Gallo, R. C. & Wong-Staal, F. *Proc. natn. Acad. Sci. U.S.A.* **83**, 772–776 (1986).
17. Schuitemaker, H. et al. *J. Virol.* **66**, 1354–1360 (1992).
18. Chesebro, B., Wehrly, K., Nishio, J. & Perryman, S. *J. Virol.* **66**, 6547–6554 (1992).
19. Greene, W. C. *New Engl. J. Med.* **324**, 308–317 (1991).
20. Peluso, R., Haase, A., Stowring, L., Edwards, M. & Ventura, P. *Virology* **147**, 231–236 (1985).
21. Haase, A. T. in *In Situ Hybridization: Applications to Neurobiology* (eds Valentino, K. L., Eberwine, J. H. & Barchas, J.D.) 197–219 (Oxford Univ. Press, New York, 1987).

ACKNOWLEDGEMENTS. We thank our colleagues P. Southern, M. Jenkins, L. Schiff and T. Reinhart for helpful discussion and C. O'Neill and T. Leonard for preparation of the manuscript and figures. We thank the NIH, the Henry M. Jackson Foundation and Bundesministerium für Forschung und Technologie for support.

## Two circadian oscillators in one cell

Till Roenneberg\* & David Morse†

\*Institut für Medizinische Psychologie, Ludwig-Maximilians-Universität, Goethestrasse 31, 8000 München 2, Germany

†Institut de Recherche en Biologie Végétale, Université de Montréal, 4101 Sherbrooke east, Montréal, Canada H1X 2B2

**A CIRCADIAN clock, which continues to oscillate in constant conditions, is almost ubiquitous in eukaryotes as well as some prokaryotes<sup>1</sup>. This class of biological oscillators drives daily rhythms as diverse as photosynthesis in plants<sup>2</sup> and the sleep-wake cycle in man<sup>3</sup> and enables organisms to anticipate environmental changes or segregate in time-incompatible processes<sup>4</sup>. Circadian oscillators share many properties, suggesting that the clock is a single mechanism, preserved throughout evolution, which is capable of controlling all the different circadian functions. Here we show that two rhythms in a unicellular organism can, under certain experimental conditions, run independently, and thus each rhythm must be controlled by its own distinct oscillator.**

The unicellular alga *Gonyaulax polyedra* displays circadian rhythms in many physiological functions, including two modes of bioluminescence<sup>5</sup> (flashing and glow) and aggregation, a type of swimming behaviour<sup>6,7</sup>. All three rhythms are highly reproducible, endogenously generated, and continue for several weeks in cultures kept in constant conditions, as shown by cells kept

in constant white light (Fig. 1a). The daily bouts of each variable appear earlier every day because the average period of each of the three rhythms is slightly shorter than 24 hours. The phase angles between the three variables are typical: flashes (grey) occur during the subjective night, the cells aggregate (striped) during the subjective day, and the glow (black) peaks at the end of the subjective night.

The individual phase relationships are more obvious when only the phases of the daily maxima are plotted (Fig. 1b). The phase difference between the flashing and the glow peaks becomes gradually smaller during the first days of the experiment, but remains stable thereafter, while the aggregation rhythm runs slightly faster than the flashing rhythm throughout the experiment and is corrected by phase jumps roughly every 7 days. Both period and phase of the glow rhythm seem to be intermediate between the two other rhythms. The *Gonyaulax* glow rhythm has both regular phase jumps<sup>8</sup> and an infradian rhythm (period  $\gg$  1 day) in its amplitude<sup>9</sup>. Although this amplitude modulation has been interpreted as the expression of a seven-day (circaseptan) clock, we propose that the glow rhythm is influenced by the relative phase angles between bioluminescence and aggregation, which are controlled by separate but coupled oscillators.

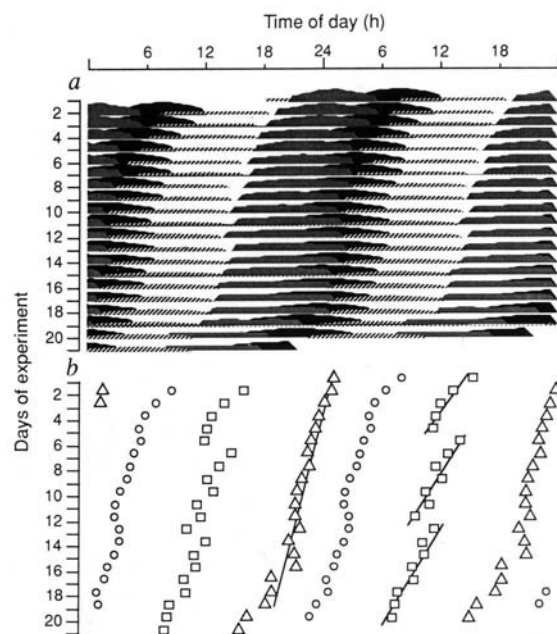


FIG. 1 Simultaneous recordings of the two bioluminescence rhythms (flashes, grey; glow, black) and the aggregation rhythm (striped) in white light ( $30 \mu\text{E m}^{-2} \text{s}^{-1}$ , 400–700 nm, Osram Universalweiss). The current setup has a moving trolley which holds a photomultiplier tube and a video camera; it can measure the luminescence and aggregation of up to 30 cultures. The methods for recording these rhythms separately have been described<sup>7,25</sup>. During the 'subjective night' (corresponding to the dark time of the entraining light cycle before the experiment, 10 p.m. to 10 a.m.), *Gonyaulax* emits bright flashes ( $100 \text{ ms}$ ,  $10^8 \text{ quanta s}^{-1}$ ) as well as a sustained glow ( $10^4 \text{ quanta s}^{-1}$ ). Luminescence is measured by a photomultiplier and digitized. The computer stores the frequency of flashes (defined by their sharp rise<sup>15</sup>) and the intensity of the glow (the averaged 16 lowest intensity values within 40 s) as separate rhythms. In the vials used for the bioluminescence measurements, cells aggregate in a dark cluster during the 'subjective day', and the computer stores the number of dark pixels in a digitized video image. In the two graphs, the vertical sequence of experimental days is doubled and replotted to the right, which allows a continuous representation of a rhythm across midnight. The non-rhythmic background is eliminated by subtracting a moving centred average of 28 h. In a, the data have been smoothed by a centred moving average of 7 h. b, The daily peak times of each rhythm (triangles, flashes; circles, glow; squares, aggregation) were determined by an algorithm using the zero transitions of the first derivative treated with a weighted moving average.



The relative independence of the bioluminescence and aggregation rhythms can be directly observed under conditions of constant dim red light (Fig. 2). The different periods of the two rhythms (flashes and aggregation) are obvious in the multiple plot of the smoothed data (Fig. 2b) while the infradian changes in amplitude, especially in the bioluminescence rhythm, can be best seen in the unsmoothed data (Fig. 2a, top curve). The differences in period are confirmed by spectral analysis of the unsmoothed data. When the entire course of the experiment is analysed, a short and a long component can be detected in each rhythm: the flashing rhythm shows a strong peak at 24.6 h and a smaller one at 22, while the aggregation rhythm shows two equally strong peaks at 24.2 and 20.6 h. If, however, 4-day windows are moved across the time series and analysed separately, the spectral analyses show single peaks, which, however, change with time (Fig. 2c). The three bouts of large period differences coincide with the crossovers of the rhythms (also indicated with arrows in Fig. 2a). Note that the transient period changes are not due to external stimuli of the experimental setup but reflect internal interactions between the two decoupled rhythms also known as relative coordination<sup>10,11</sup>. The relative amounts of these phase-dependent period changes (Fig. 2c) indicate that the aggregation rhythm is more strongly influenced by the bioluminescence rhythm than vice versa.

Because two circadian rhythms may adopt new relative phase relationships when released to constant conditions, or when

conditions are changed, their periods may differ until the new phase relationship is stabilized. Smaller period differences (~1 h) have been reported for glow and flashing<sup>12,13</sup>, but in those records it was difficult to distinguish between a reestablishment of a new relative phase relationship and a true internal desynchronization of two rhythms. In our red light experiments, the two rhythms cross each other several times, which excludes the possibility that they are merely adopting a new phase relationship.

We were curious as to why, in spite of extensive studies of the circadian system of *Gonyaulax*, internal desynchronization has not been previously noted<sup>14</sup>, and it seems likely that some critical factor was supplied by our experimental setup. Unlike recordings of the aggregation rhythm alone, cells are exposed to darkness while luminescence is measured, which shortens the glow rhythm by about 1 h when the cells are kept in constant white light<sup>15</sup>. However, when cells are exposed, for example, to a 45-min: 15-min red light: dark cycle, the period of the aggregation rhythm is shortened by almost 2 h (Fig. 3b), whereas there is no effect on the period of the flashing rhythm (Fig. 3a). Although internal desynchronization is only obvious when the periods differ significantly, period differences as such and a change in coupling strength are two different aspects of the phenomenon. Dim red light may be responsible for the reduction of the coupling strength while the frequent dark pulses make the internal desynchronization apparent by exclusively accelerating the aggregation rhythm.

Internal desynchronization as such is not a new phenomenon, and has also been shown for the temperature rhythm and the sleep-wake cycle in humans<sup>16,17</sup>. As in *Gonyaulax*, the two rhythms show relative coordination in their decoupled state.

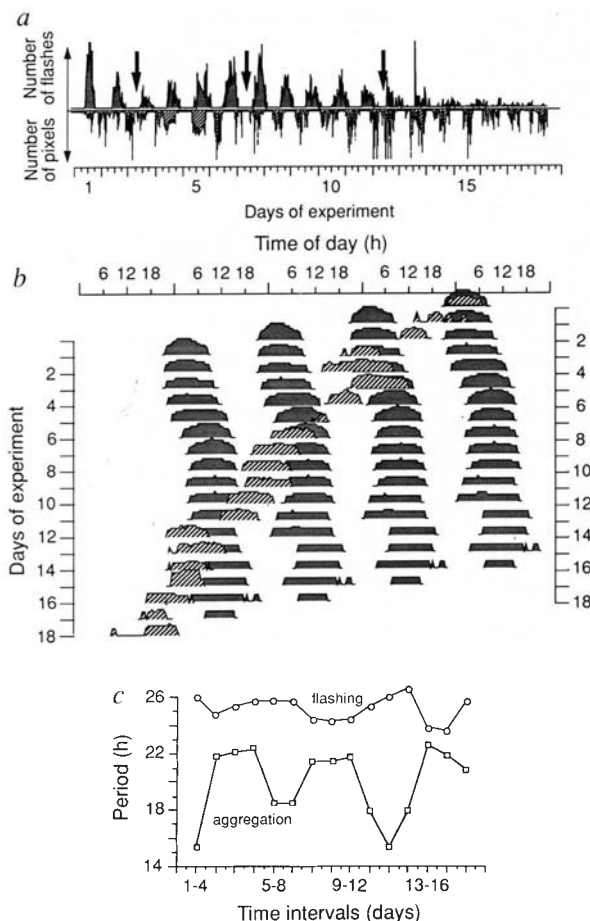


FIG. 2 Bioluminescence and aggregation in constant red light ( $13 \mu\text{E m}^{-2} \text{s}^{-1}$ , filter: Rosco Lux number 19,  $>600 \text{ nm}$ ,  $21^\circ\text{C}$ ). a, The unsmoothed data; the three crossovers of the two rhythms are indicated by arrows. For the multiple plot (b), the data have been treated as described for Fig. 1. The complex period changes during the course of the experiment are analysed quantitatively by calculating the spectral analysis<sup>26</sup> for 4-day windows, which are moved across the time series of the unsmoothed data (c).

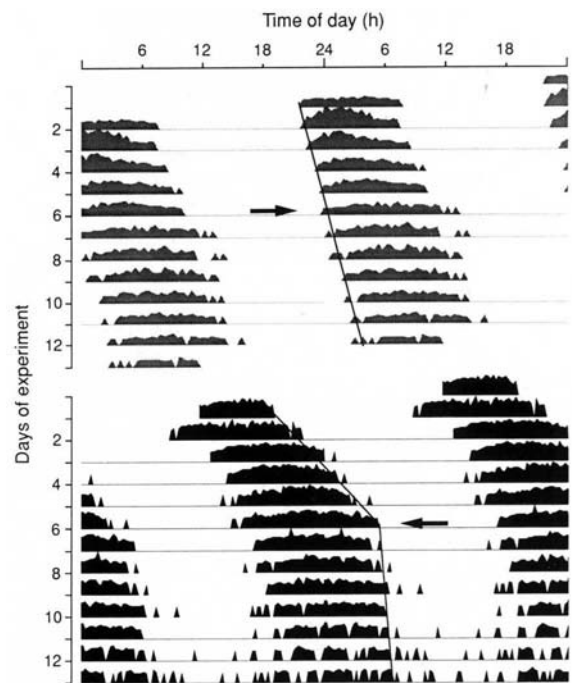


FIG. 3 The effect of frequent dark pulses on the flashing and aggregation rhythm were investigated by recording the two rhythms separately but under identical conditions (red light:  $13 \mu\text{E m}^{-2} \text{s}^{-1}$ ,  $22^\circ\text{C}$ ). For 5 days, cells in the aggregation setup were exposed to true constant light while those in the bioluminescence setup experienced the usual frequent dark pulses (4 min every 25 min) during the luminescence measurements. Note that the period of the aggregation rhythm (lower graph) is longer than that of the flashing rhythm (top graph). On day six, frequent dark pulses (15 min every hour) were introduced in both setups for the rest of the experiment. Although the flashing rhythm is unaffected, the aggregation rhythm is shortened by almost 2 h.

Whether these separate rhythmic processes in humans are actually driven by separate self-sustained oscillators with pacemaker properties is still being discussed<sup>18</sup>. In any event, the presence of different oscillators in multicellular organisms is not conceptually difficult, as they could be localized in different cell types or tissues. This is not the case in our experiments, as all cells undergo both rhythms.

The demonstration of two discrete timing mechanisms in a single cell organism is of major importance to our understanding of the nature and origin of biological clocks. Instead of one clock controlling all circadian rhythms of the cell, there are at least two, and at the limit, a separate oscillator for each rhythm (a possibility proposed for the *Gonyaulax* bioluminescence system more than 30 years ago<sup>19</sup>). Although it is too early to speculate on the nature or subcellular location of the oscillatory

processes regulating bioluminescence and aggregation, it seems reasonable that aggregation, which parallels vertical migration<sup>7</sup>, is related to photosynthesis, as this would be more efficient at the surface. The involvement of the chloroplast in circadian phenomena has often been proposed<sup>20</sup>, and experiments in the giant alga *Acetabularia* have suggested that the nucleus and the chloroplast contribute differently to the circadian system<sup>21,22</sup>. Because cyanobacteria also possess a circadian system<sup>4</sup> and may be related to the evolutionary ancestors of photosynthetic organelles<sup>23,24</sup>, it is possible that chloroplasts have contributed their own oscillator to the circadian systems of the eukaryotic cell. It will thus be of great interest to determine the total number of oscillators in *Gonyaulax* as well as in other algal species that lack chloroplasts. □

Received 8 June 1992; accepted 14 January 1993.

- Hastings, J. W., Rusak, B. & Boulos, Z. In *Neural and Integrative Animal Physiology* (ed. Prosser, C. L.) 435–546 (Wiley-Liss, Cambridge, 1991).
- Hastings, J. W., Astrachan, L. & Sweeney, B. M. *J. gen. Physiol.* **45**, 69–76 (1961).
- Wever, R. *The Circadian System of Man* (Springer, Berlin, 1979).
- Huang, T.-C., Tu, J., Chow, T.-L. & Chen, T.-H. *Pl. Physiol.* **92**, 531–533 (1990).
- Hastings, J. W. & Sweeney, B. M. *Biol. Bull.* **115**, 440–458 (1958).
- Roenneberg, T., Colfax, G. N. & Hastings, J. W. *J. biol. Rhythms* **4**, 201–216 (1989).
- Roenneberg, T. & Hastings, J. W. In *Oscillations and Morphogenesis* (ed. Rensing, L.) 399–412 (Dekker, New York, 1992).
- Morse, D., Fritz, L. & Hastings, J. W. *Trends biochem. Sci.* **15**, 262–265 (1990).
- Halberg, F., Cornelissen, G. & Broda, H. *Cell Biophys.* **8**, 69–85 (1985).
- Holst, E. *Ergebn. Physiol.* **42**, 228–306 (1939).
- Aschoff, J. In *Biological Rhythms* (ed. Aschoff, J.) 81–93 (Plenum, New York, 1981).
- Krasnow, R. et al. *J. comp. Physiol.* **138**, 19–26 (1980).
- von der Heyde, F., Wilkens, A. & Rensing, L. *J. biol. Rhythms* **7**, 115–124 (1992).
- McMurry, L. & Hastings, J. W. *Science* **175**, 1137–1139 (1972).
- Broda, H., Gooch, V. D., Taylor, W. R., Aiuto, N. & Hastings, J. W. *J. biol. Rhythms* **1**, 251–263 (1986).

- Aschoff, J. *Science* **148**, 1427–1432 (1965).
- Aschoff, J., Gerecht, U. & Wever, R. *Jap. J. Physiol.* **17**, 450–457 (1967).
- Daan, S., Beersma, D. G. M. & Borbély, A. A. *Am. J. Physiol.* **246**, R161–R178 (1984).
- Hastings, J. W. & Sweeney, B. M. *Proc. natn. Acad. Sci. U.S.A.* **43**, 804–811 (1957).
- Edmunds, L. N. Jr *Cellular and Molecular Bases of Biological Clocks: Models and Mechanisms of Circadian Time Keeping* (Springer, New York, 1988).
- Schweiger, E., Wallraff, H. G. & Schweiger, H. G. *Science* **146**, 658–659 (1964).
- Schweiger, E., Wallraff, H. G. & Schweiger, H. G. *Z. Naturforsch.* **19c**, 499–505 (1964).
- Schwartz, R. M. & Dayhoff, M. D. *Science* **199**, 395–403 (1978).
- Margulis, L. *Symbiosis in Cell Evolution* (Freeman, New York, 1991).
- Taylor, W., Wilson, S., Presswood, R. P. & Hastings, J. W. *J. Interdiscipl. Cycle Res.* **13**, 71–79 (1982).
- Box, G. E. P. & Jenkins, G. M. *Time Series Analysis—Forecasting and Control* (Holden-Day, San Francisco, 1970).

ACKNOWLEDGEMENTS. We thank J. W. Hastings for his continuous input to our work and I. Neher for technical assistance. This paper is dedicated to J. Aschoff. This work was supported by grants from the Deutsche Forschungsgemeinschaft and the National Science and Engineering Research Council of Canada.

## Mechanical basis of meiotic metaphase arrest

Kim S. McKim\*, Janet Ko Jang\*,  
William E. Theurkauf† & R. Scott Hawley\*‡

\* Department of Genetics, University of California, Davis, California 95616, USA

† Department of Biochemistry and Cell Biology, SUNY Stonybrook, Stonybrook, New York 11794-5215, USA

**CONTROL** of the metaphase to anaphase transition is a central component of cell-cycle regulation. Arrest at either metaphase I or II before fertilization is a common component of oogenesis in many organisms<sup>1</sup>. In *Drosophila melanogaster* females, this arrest occurs at meiosis I with the chiasmate bivalents tightly massed at the metaphase plate and the nonexchange chromosomes positioned between the plate and the poles on long tapered spindles<sup>2</sup>. Meiosis resumes only after passage through the oviduct<sup>3,4</sup>. Thus, metaphase arrest defines an important checkpoint in the meiotic cell cycle<sup>1</sup>. We report here that this arrest results from the balancing of chiasmate bivalents at the metaphase plate. Two meiotic mutations, *mei-9<sup>b</sup>* and *mei-218<sup>a4</sup>*, both of which greatly reduce the frequency of chiasma formation, bypass the metaphase block and allow stage 14 oocytes to finish both meiotic divisions without arrest. We conclude that metaphase arrest results from the balancing of kinetochore forces due to chiasmata.

Recombination is reduced in *mei-9<sup>b</sup>* and *mei-218<sup>a4</sup>* mutant females to about 16% and 8% of control levels, respectively<sup>5–7</sup>. Our confocal microscopic analysis of oocytes from wild-type females and from females homozygous for *mei-9<sup>b</sup>* or *mei-218<sup>a4</sup>* revealed no obvious differences in the structure of prometaphase figures (Fig. 1a). But oocytes from females homozygous for *mei-9<sup>b</sup>* or *mei-218<sup>a4</sup>* did not display the accumulation of meta-

phase I figures characteristic of oocytes from wild-type females (Fig. 1b; Table 1). Rather, there were large numbers of anaphase I figures (Fig. 1d–g).

This progression into anaphase I is heralded by the movement of all chromosomes away from the metaphase plate to the poles. Moreover, the nonexchange chromosomes, such as the 4th, which are positioned between the plate and the poles during metaphase arrest, now complete their poleward migration. Therefore, the movement of chromosomes towards the poles is coupled with a release of the plateward force which normally opposes poleward migration<sup>2,8</sup>. Meiosis then proceeds until

FIG. 1 Microtubules (red) and chromatin (yellow-green) in wild-type and mutant oocytes showing prometaphase (a, *mei-218<sup>a4</sup>/FM7*), metaphase I (b, *mei-218<sup>a4</sup>/FM7*; c, *mei-9<sup>b</sup>/mei-9<sup>b</sup>*), early anaphase I (d, *mei-9<sup>b</sup>/mei-9<sup>b</sup>*), anaphase I (e–g, *mei-218<sup>a4</sup>/mei-218<sup>a4</sup>*), metaphase II (h, *mei-218<sup>a4</sup>*) and anaphase II (i, *mei-218<sup>a4</sup>/mei-218<sup>a4</sup>*). In *Drosophila* females, the meiotic spindle assembles during stages 12 to 14 of oogenesis<sup>2,4</sup> (prometaphase to metaphase arrest). During spindle assembly and elongation at prometaphase, the achiasmate chromosomes precociously move towards the pole while the chiasmate chromosomes remain at the metaphase plate<sup>2,14</sup>. Meiosis normally arrests in wild-type oocytes at metaphase I (refs 2–4). METHODS. *FM7* is a multiply inverted X chromosome and carries wild-type alleles of both *mei-9* and *mei-218* (ref. 11). Oocytes were prepared and examined as described previously<sup>2</sup>. This process is briefly summarized as follows. Egg chambers were collected from 3–6-day-old females. To extract the chambers, the females were collected into modified Robb's medium and broken apart by quick pulses of a blender. Late-stage oocytes were collected by passage of the mixture through a loose mesh. The oocytes were fixed for 5–10 min on a rotator at room temperature in a hypertonic solution, the latter assuring that the mature oocytes were not hypotonically activated. After removal of the follicle cells, chorion and vitelline membranes, the oocytes were permeabilized with 1% Triton X-100 in PBS, and then double labelled with anti-histone antibody conjugated to fluorescein (yellow-green fluorescence) and anti-tubulin conjugated to rhodamine (red fluorescence).

‡ To whom correspondence should be addressed.



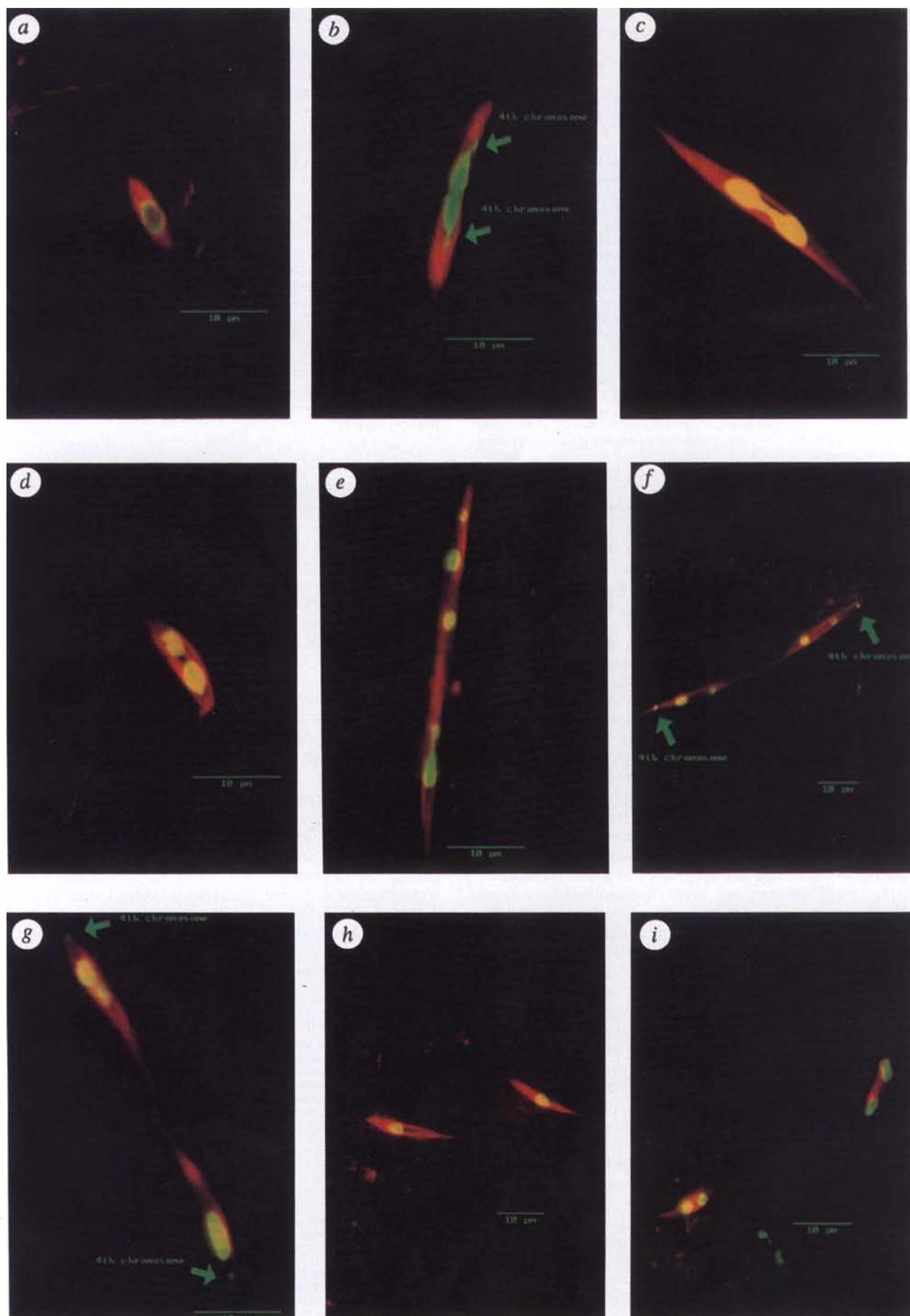


TABLE 1 Classification of wild-type and mutant oocyte nuclei

| Genotype                    | Prometaphase | Metaphase | Meiotic stage |              |             |
|-----------------------------|--------------|-----------|---------------|--------------|-------------|
|                             |              |           | Anaphase I    | Metaphase II | Anaphase II |
| Wild-type                   | 17           | 30 (30)   | 0             | 0            | 0           |
| <i>mei-9<sup>b</sup></i>    | 21           | 31 (33)   | 12            | 8            | 3           |
| <i>mei-218<sup>a4</sup></i> | 29           | 15 (16)   | 18            | 14           | 9           |

Meiotic stages were defined as follows. Prometaphase, the karyosome is surrounded either by diffuse spindles or by short untapered spindles; metaphase I, elongated and tapered spindles surrounding stretched chromatin masses at the metaphase plate; anaphase I, chromosome masses are clearly separated and positioned between the centre of the spindle and the poles; metaphase II, two separate spindles each encompassing a single chromatin mass; anaphase II, two separate spindles each with two chromosome masses that are clearly separated and positioned between the centre of the spindle and the poles. It is clear that we are observing true meiosis II figures, rather than a repeat of metaphase-anaphase I figures because in the metaphase II and anaphase II figures the nonexchange chromosomes do not move precociously to the poles (Fig. 1*h, i*), as is characteristic of metaphase I and anaphase I figures. The expected number of metaphases is listed in parentheses and was calculated by multiplying the total number of metaphase I, anaphase I, metaphase II and anaphase II figures by the fraction of oocytes expected to have at least one chiasma ( $1 - (E_0)^5$ ). In the case of *mei-218<sup>a4</sup>*,  $E_0$  frequencies were available for both the X chromosome (0.95) and for the left arm of chromosome 2 (0.93) (refs 5–7). Accordingly, the fraction of oocytes expected to have at least one chiasma was calculated as  $1 - (0.95)(0.93)^4$ . In the case of *mei-9<sup>b</sup>*,  $E_0$  frequencies were available only for the left arm of chromosome 2 (0.83) (refs 5–7). In this case the fraction of oocytes expected to have at least one chiasma was calculated as  $1 - (0.83)^5$ .

completion as indicated by the appearance of metaphase II (Fig. 1*h*) and anaphase II (Fig. 1*i*) figures.

It is clear that we are observing true meiosis II figures, rather than a repeat of metaphase-anaphase I figures because in the metaphase II and anaphase II figures the small 4th chromosomes do not move precociously to the poles (Fig. 1*h, i*). Rather, the paired spindles in the mutant oocytes have all of the chromosomes massed at the metaphase plate (Fig. 1*g*), which is the normal metaphase II configuration. Thus, in homozygous *mei-9<sup>b</sup>* or *mei-218<sup>a4</sup>* mutant females, most oocytes do not arrest at metaphase I. Instead, they complete the meiotic cycle.

Because *mei-9<sup>b</sup>* and *mei-218<sup>a4</sup>* do not completely suppress chiasma formation, some fraction of oocytes should possess at least one chiasma, causing arrest at metaphase I. Metaphase figures identical to those observed in wild-type females, with chiasma visible between the two chromatin masses, are occasionally observed in both *mei-9<sup>b</sup>* and *mei-218<sup>a4</sup>* females (Fig. 1*b, c*). The observed frequency of metaphase figures is very close to the calculated frequency of meiosis in which one or more exchanges occur (indicated in parentheses in Table 1). In addition, metaphase I oocytes are more frequently found from females homozygous for *mei-9<sup>b</sup>* than from females homozygous for *mei-218<sup>a4</sup>*. This is consistent with the higher frequency of recombination in *mei-9<sup>b</sup>* females<sup>5,6</sup>. These correlations suggest that even one chiasma is sufficient to ensure metaphase arrest.

The precocious completion of the meiotic cycle does not appear to cause a significant degree of zygotic lethality. As shown in Table 2, the fraction of viable zygotes produced by *mei-218<sup>a4</sup>* females (25%) is no greater than would be expected based on the inviability of aneuploid zygotes receiving zero or two maternal copies of an X chromosome or a major autosome (25%).

The finding that essentially all of the zygotic mortality in these mutants can be accounted for by the nondisjunction of achiasmate autosomes further demonstrates that the lack of chiasmata results in an otherwise normal passage through meiosis. If, for example, the 'meiosis II' spindles we observe in the mutant oocytes were actually two aberrant meiosis I spindles, then the ensuing four meiosis II divisions would result in products that are grossly deficient in chromosome content. Thus, we would predict that activation and fertilization of these oocytes would produce inviable progeny at a much higher frequency than predicted from the random disjunction of achiasmate bivalents alone.

TABLE 2 Zygotic viability of progeny from wild-type and *mei-218<sup>a4</sup>*

| Maternal genotype                                | Eggs  | Larvae | Survival | Expected |
|--------------------------------------------------|-------|--------|----------|----------|
| <i>mei-218<sup>a4</sup>/FM7</i>                  | 720   | 691    | 0.96     | 0.99     |
| <i>mei-218<sup>a4</sup>/mei-218<sup>a4</sup></i> | 1,001 | 254    | 0.25     | 0.25     |

Owing to the paucity of meiotic recombination, females homozygous for *mei-218<sup>a4</sup>* display high levels of meiotic nondisjunction affecting achiasmate homologues at high frequencies<sup>7,13</sup>. Fertilization of ova which carry zero or two maternal copies of a major autosome will be zygotic lethals. Similarly, one half of the ova bearing zero or two X chromosomes will produce inviable zygotes. Although the situation is considerably more complex with respect to nonhomologous segregation<sup>5–7</sup>, for our purposes it is correct to calculate the expected frequency of inviable zygotes that results from aneuploid ova as follows. When both X chromosomes and a pair of major autosomes are nonexchange, or when two pairs of autosomes are nonexchange, each chromosome segregates independently, or very nearly so, of their homologues<sup>7,13</sup>. As most pairs of chromosomes are nonexchange in *mei-218<sup>a4</sup>* females, the frequency of nondisjunction for the two major autosomes should be about  $1/2(E_0)^2$  or 43.2% and for the X-chromosome should be  $\sim 1/4(E_0)$  or 24%. Thus, the frequency of viable zygotes can be calculated as  $(1 - 0.24)(1 - 0.43)(1 - 0.43) = 0.25$ . These estimates are highly concordant with those obtained by Carpenter using a different methodology<sup>10</sup>.

*mei-9* and *mei-218* are required for different aspects of meiotic recombination. *mei-9* is required late in the recombination pathway, possibly in the resolution of heteroduplexes<sup>9</sup>, and forms morphologically normal recombination nodules<sup>10</sup>. In contrast, *mei-218* is required earlier, possibly for the transition from gene conversion to reciprocal exchange<sup>9</sup>, and has morphologically abnormal recombination nodules<sup>10</sup>. Thus, these two mutants fail to form chiasma for different reasons, suggesting it is the lack of chiasma formation, and not some pleiotropic effect on the meiotic process, which prevents metaphase arrest. This conclusion is further supported by our observation that a third recombination-defective mutation, *mei-W68* (ref. 11), also prevents metaphase arrest (data not shown).

One chiasmate bivalent is both necessary and sufficient to confer metaphase arrest. Although metaphase arrest can thus be explained solely in terms of the physical ability of chiasmata to balance kinetochore forces and hold bivalents at the metaphase plate<sup>12</sup>, our data do not preclude the existence of biochemical regulatory systems that sense mechanical tension at the plate. Preventing chiasma formation eliminates the arrest and allows the meiotic cycle to proceed. The fact that meiosis is then completed demonstrates that there are no further arrest points in the meiotic process.

Perhaps it is not surprising that in a meiotic system like that in *Drosophila*, where the chromosomes organize the meiotic spindle<sup>2</sup>, the chromosomes themselves control the cell cycle.

Received 6 November 1992; accepted 15 January 1993.

- Murray, A. W. *Nature* **359**, 599–604 (1992).
- Theurkauf, W. E. & Hawley, R. S. *J. Cell Biol.* **116**, 1167–1180 (1992).
- Sonnenblick, B. P. in *Biology of Drosophila* (ed. Demerec, M.) 62–167 (Hafner, New York, 1965).
- King, R. C. *Ovarian Development in Drosophila melanogaster* (Academic, New York, 1970).
- Carpenter, A. T. C. & Baker, B. S. *Genetics* **71**, 255–286 (1972).
- Carpenter, A. T. C. & Sandler, L. *Genetics* **76**, 453–475 (1974).
- Baker, B. S. & Hall, J. C. in *The Genetics and Biology of Drosophila* Vol. 1a (eds Ashburner, M. & Novitski, E.) 351–434 (Academic, New York, 1976).
- Zhang, P. & Hawley, R. S. *Genetics* **125**, 115–127 (1990).
- Carpenter, A. T. C. *Cold Spring Harb. Symp. quant. Biol.* **49**, 23–29 (1984).
- Carpenter, A. T. C. *Chromosoma* **75**, 259–292 (1979).
- Lindsley, D. L. & Zimm, G. G. *The Genome of Drosophila melanogaster* (Academic, San Diego, 1992).
- Nicklas, R. B. *Genetics* **78**, 205–213 (1974).
- Hawley, R. S. in *Genetic Recombination* (eds Kucherlapati, R. & Smith, G. R.) 497–527 (American Society for Microbiology, Washington DC, 1988).
- Puro, J. & Noll, K. *Chromosoma* **63**, 273–286 (1977).

ACKNOWLEDGEMENTS. We thank T. Arbel for critical insights into the cytological analysis. We also thank K. Burtis, A. T. C. Carpenter, C. Holm, H. Irick, N. Kleckner, T. Orr-Weaver, J. Scholey and E. Shuster for discussions and comments on the manuscript. This work was supported by a Natural Sciences and Engineering Research Council (Canada) fellowship to K.S.McK., an American Cancer Society Faculty Research award to R.S.H., and by grants from the National Science Foundation and the American Cancer Society to R.S.H.





The molecule and variants with mutations in either the loop or framework residues were produced by solid-phase peptide synthesis<sup>8,16</sup>.

The minibody is a monomer, as assessed by size-exclusion chromatography on a Sephadex G-50 column calibrated with a series of compact globular proteins as molecular size markers (data not shown). The apparent  $M_r$  (7.2K) agrees well with the  $M_r$  of the monomer calculated from the amino-acid sequence (7.112K). Identical chromatograms were obtained at pH 5.8 and 7.0, and when a large molar excess of  $\text{ZnCl}_2$  (0.1 mM) was included in the eluent.

The solubility is about 10  $\mu\text{M}$ , below the millimolar range required for nuclear magnetic resonance or X-ray crystallographic structural studies, but suitable for analysis by ultraviolet circular dichroism (CD) spectroscopy. The CD spectrum shows characteristic  $\beta$ -structure features and the estimated secondary structure content agrees well with our model (Fig. 2a). The stability of the molecule to denaturation was measured by monitoring the mean residue ellipticity at 217 nm as a function of urea concentration (Fig. 2b). The curve shows a rather sharp cooperative transition up to 4.0 M urea with a midpoint at 2.15 M, followed by a slower decrease in ellipticity at increasing denaturant concentration, which does not reach a plateau up to 9.0 M urea. Goto *et al.*<sup>17</sup> obtained the same profile for the monomeric form of the constant domain CL of a  $\lambda$  Bence-Jones protein, which is structurally similar to the immunoglobulin VH domain and thus to the minibody model. They estimated a free energy of denaturation, extrapolated at zero concentration of denaturant,  $\Delta G_D^{\text{H}_2\text{O}}$  of 6 kcal mol<sup>-1</sup> for the intact CL and 1.8 kcal mol<sup>-1</sup> for the same domain with the single disulphide bond reduced, and attributed most of this difference in stability to the increased entropy of the reduced CL fragment in the denatured state. For the unfolding transition of the minibody we estimated<sup>18</sup> a  $\Delta G_D^{\text{H}_2\text{O}}$  of 2.5 kcal mol<sup>-1</sup>, a value close to that found for the reduced CL. So far, no disulphide bond has been engineered in the minibody.

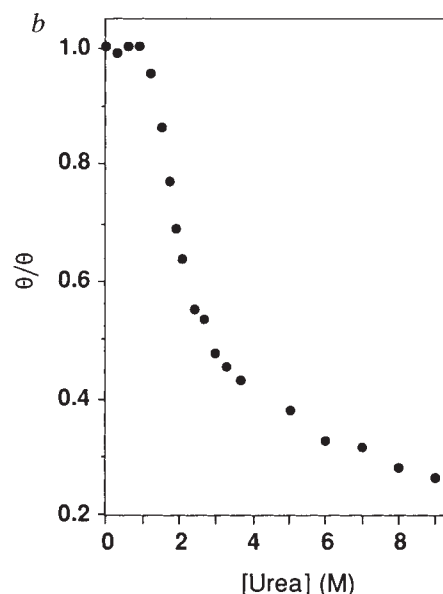
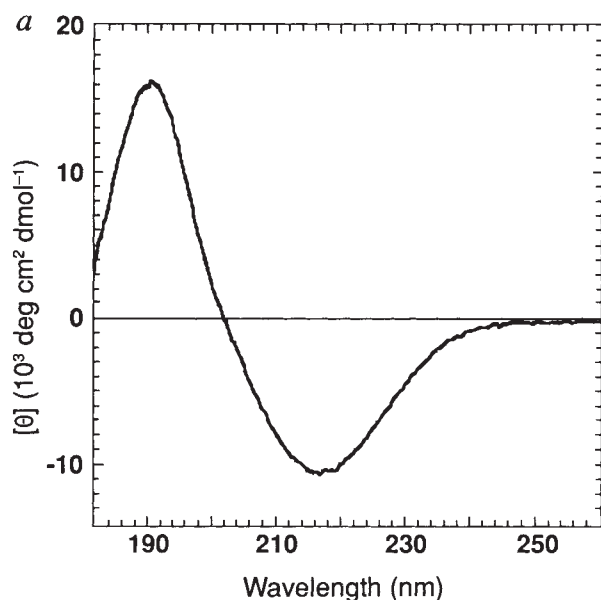


FIG. 2 Analysis of minibody structure by CD. *a*, CD spectra were obtained at 21 °C in 10 mM sodium phosphate buffer, pH 5.8 and in 10 mM MOPS at pH 7.0, in the presence and absence of 1 mM  $\text{Zn(II)}$ . The spectra are superimposable in the region (190–250 nm) where the MOPS absorbance does not interfere (MOPS was selected because of the lower solubility of the molecule in phosphate at pH 7.0). We show here the 182–260 nm spectrum obtained at pH 5.8. The estimated secondary structure content<sup>24</sup> is 0%  $\alpha$ -helix, 65%  $\beta$ -structure, 31% turn and 3% unordered, well in agreement with the values calculated for the model<sup>25</sup>, which are 0%  $\alpha$ -helix, 72%  $\beta$ -structure, 18%

As predicted, the minibody molecule binds metals. A lower limit for the dissociation constant ( $K_d$ ) of the  $\text{Zn(II)}$ -minibody complex (about  $10^{-6}$  M) was estimated by a filter binding assay (data not shown). The metal-binding activity was characterized by a chromatographic  $^{65}\text{Zn(II)}$  binding assay and by filter-binding competition experiments. We proved that: (1) histidines, assumed to have normal  $pK_a$  values, are responsible for this activity (Fig. 3a); (2) three histidines with the same linear distance in the sequence but in random coil conformation do not confer activity on the molecule (Fig. 3b); (3) in the minibody, both loops contribute to the metal-binding site, because the His13Ala mutant<sup>8</sup>, in which the chelating histidine at position 13 in the H1 loop is replaced by alanine, shows a fourfold decrease in affinity (data not shown); the CD spectrum of this mutant is superimposable on that of the minibody; (4) the binding site is selective for different ions, with an order of affinities mirroring the one observed for carbonic anhydrase, namely,  $\text{Cu} > \text{Zn} \gg \text{Cd} > \text{Co}$  (ref. 19; Fig. 3c). Notably,  $\text{Ca}^{2+}$  and  $\text{Mg}^{2+}$  do not inhibit metal binding.

All our results show that the design was successful. Both the metal-free and metal-bound forms of the minibody are compact globular monomers in solution. The secondary structure content agrees well with the expected one, the molecule shows a sharp denaturant-induced unfolding transition, and its  $\Delta G_D^{\text{H}_2\text{O}}$  value, although lower than the values observed for natural proteins (5–15 kcal mol<sup>-1</sup>)<sup>20</sup>, is very similar to that of the aforementioned reduced immunoglobulin CL domain. The metal-binding site has a reasonably good affinity, comparable to that of other engineered three-histidine binding sites<sup>13</sup>. Its  $K_d$  value is intermediate between those observed for naturally occurring two-histidine<sup>14</sup> and three-histidine binding sites<sup>15</sup>. We have shown that at least one histidine from each loop participates in metal binding, which indicates that the loops are close in space as predicted by our model. The limited solubility of the molecule does not allow a more detailed characterization of the structure of the metal-binding site. Therefore, we are now directing our

turn and 10% unordered. *b*, Urea denaturation curve.  $\theta/\theta_0$  is the ratio of ellipticity at 217 nm at the indicated molarity of urea to the ellipticity at zero concentration of denaturant.

**METHODS.** Stock solutions were prepared in 1,1,1-3,3,3-hexafluoro-2-propanol; aliquots of these solutions were diluted 100-fold with the buffer (with or without urea) to a final concentration of 0.08 mg ml<sup>-1</sup> (determined by quantitative amino-acid analysis). Spectra were taken on a Jasco J-710 spectropolarimeter in a 0.1-cm-pathlength cell, with a 16-s time constant and a 2 nm min<sup>-1</sup> scan speed.



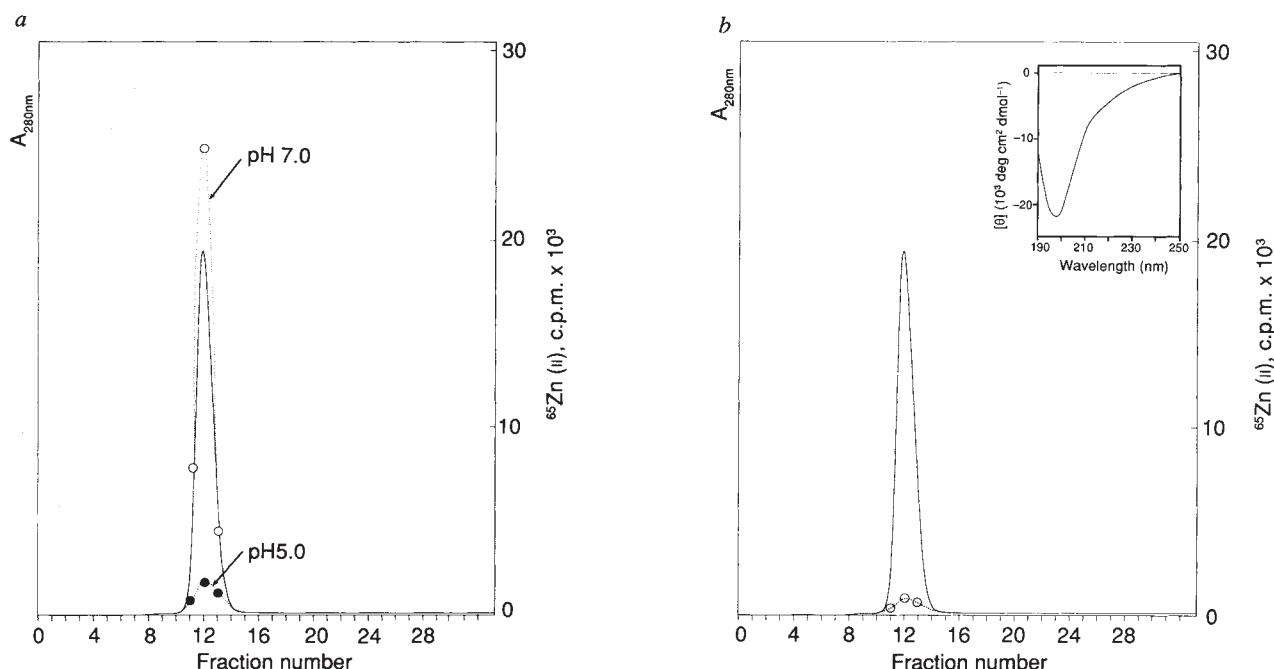
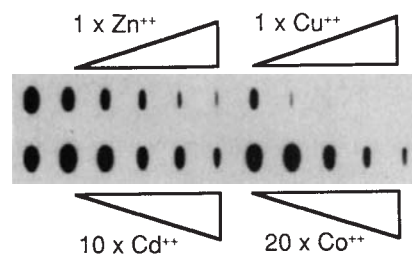


FIG. 3 Zinc-binding activity of the minibody and one mutant. The activity was assayed by binding to <sup>65</sup>Zn(II) in a gel-filtration assay. Solid lines: absorbance at 280 nm (A<sub>280</sub>); dotted lines: <sup>65</sup>Zn(II) c.p.m. a, At pH 7.0, the minibody and the <sup>65</sup>Zn(II) coelute from the column. No significant binding is observed at pH 5.0, when histidine residues are protonated. b, We synthesized a mutant with 10 amino-acid substitutions in the framework (Arg1Ala, Gln4Pro, Glu17Asp, Val19Gly, Arg20Lys, Arg29Lys, Lys34Asp, Glu38Asp, Arg46Lys, Ile48Ser) but retaining the three histidines in the same position in the sequence<sup>16</sup>. This mutant, whose CD spectrum shows a characteristic random coil conformation (see inset), does not bind the metal. c, Filter binding competition assay, which shows that the metal-binding site is selective for different ions.

METHODS. a, b, 5 nmol of each peptide was incubated for 30 min with 20 μCi of <sup>65</sup>Zn(II) (1,000 mCi mg<sup>-1</sup> zinc) in 10 mM MOPS, pH 7.0, and chromatographed on a Sephadex G-15 column (1 × 17 cm). The collected fractions (1 ml) were counted in a β-counter. Under these conditions, free <sup>65</sup>Zn(II) was slowly eluted beyond the total column volume. c, Peptide (200 pmol) was incubated at room temperature for 20 min in 100 μl binding buffer (200 mM



MOPS pH 7.0, 10 mM CaCl<sub>2</sub>, 10 mM MgCl<sub>2</sub>, 2 mM CHAPS) with 0.35 μCi of <sup>65</sup>Zn (1,000 mCi mg<sup>-1</sup>), in the presence of increasing amounts of unlabelled metal (ZnCl<sub>2</sub> and CuCl<sub>2</sub>: 1–5 nmol; CdCl<sub>2</sub>: 10–50 nmol; CoCl<sub>2</sub>: 20–100 nmol). After filtering through 0.22-μm nitrocellulose filters and washing with the binding buffer, filters were autoradiographed. The first slot in each lane is without competitor.

efforts to the improvement of solubility by further engineering of the minibody framework.

The His13Ala mutant and other mutants in the loops all show CD spectra superimposable with that of the minibody (E.B. *et al.*, manuscript in preparation). Furthermore, in the design we quite liberally modified the sequences of the template loops to introduce the metal-chelating residues. This suggests that the minibody loops may be tolerant to sequence changes, as are their counterparts in immunoglobulins.

This has interesting practical applications: we have genetically constructed a conformationally constrained peptide library by

randomly mutagenizing the loops of the minibody, fused to the pIII protein of filamentous phage f1 for display<sup>21</sup> (M.S. *et al.*, manuscript in preparation). The advantage of using the minibody is threefold. First, this library displays discontinuous conformational epitopes. Second, we believe that the existing knowledge on the structurally related immunoglobulin loops may simplify the prediction of the conformation of the binding loops selected from the library, a key step to design peptidomimetic derivatives<sup>22</sup>. Third, the small size of the molecule might help the experimental determination of the structure of the putative complexes. □

Received 23 November 1992; accepted 3 February 1993.

1. Chothia, C. & Lesk, A. M. *J. molec. Biol.* **196**, 901–917 (1987).
2. Brucoleri, R. E., Haber, E. & Novotny, J. *Nature* **335**, 564–568 (1988).
3. Martin, A. C. R., Cheetham, J. C. & Rees, A. R. *Meth. Enzym.* **203**, 121–153 (1991).
4. Richardson, J. S. & Richardson, D. C. *Trends biochem. Sci.* **14**, 304–309 (1989).
5. Hill, C. P., Anderson, D. H., Wesson, L., DeGrado, W. F. & Eisenberg, D. *Science* **249**, 543–546 (1990).
6. Hecht, M. H., Richardson, J. S., Richardson, D. C. & Odgen, R. C. *Science* **249**, 884–891 (1990).
7. Fedorov, A. N. *et al. J. molec. Biol.* **225**, 927–931 (1992).
8. Bianchi, E., Sollazzo, M., Tramontano, A. & Pessi, A. *Int. J. Peptide Protein Res.* (in the press).
9. Kubiak, T., Whitney, D. B. & Merrifield, R. B. *Biochemistry* **26**, 7849–7855 (1987).
10. Segal, D. M. *et al. Proc. natn. Acad. U.S.A.* **71**, 4298–4302 (1974).
11. Regan, L. & Clarke, N. D. *Biochemistry* **29**, 10878–10883 (1990).
12. Higaki, J. N., Haymore, B. L., Chen, S., Fletterick, R. J. & Craik, C. S. *Biochemistry* **29**, 8582–8586 (1990).
13. Iverson, B. L. *et al. Science* **249**, 659–662 (1990).

14. Arnold, F. H. & Haymore, B. L. *Science* **25**, 1796–1797 (1991).
15. Vallee, B. L. & Auld, D. S. *Biochemistry* **29**, 5647–5656 (1990).
16. Bianchi, E., Sollazzo, M., Tramontano, A. & Pessi, A. *Int. J. Peptide Protein Res.* (in the press).
17. Goto, Y. & Hamaguchi, K. *J. Biochem.* **86**, 1433–1441 (1979).
18. Pace, C. N. *Meth. Enzym.* **131**, 266–280 (1986).
19. Lindskog, S. & Nyman, P. O. *Biochim. biophys. Acta* **85**, 462–474 (1964).
20. Creighton, T. E. *Biochem. J.* **270**, 1–16 (1990).
21. Smith, G. P. *Curr. Opin. Biotechnol.* **2**, 668–673 (1991).
22. Saragovi, H. U., Greene, M. I., Chrusciel, R. A. & Kahn, M. *Biotechnology* **10**, 773–778 (1992).
23. Kannan, K. K. *et al. Proc. natn. Acad. Sci. U.S.A.* **72**, 51–55 (1975).
24. Chang, X. T., Wu, C. S. C. & Yang, J. Y. *Analyt. Biochem.* **91**, 13–31 (1978).
25. Levitt, M. & Greer, J. *J. molec. Biol.* **114**, 181–293 (1977).

ACKNOWLEDGEMENTS. We thank R. Cortese for suggesting two critical experiments, A. M. Lesk for advice and encouragement, R. De Francesco, A. Wallace, J. Jiricny, G. Ciliberto and A. S. Verdini for critical reading of the manuscript, M. Carson for the program RIBBONS, and Y. M. Cully for the artwork.

## Identification of an Alu retrotransposition event in close proximity to a strong candidate gene for Huntington's disease

Y. Paul Goldberg\*, Johanna M. Rommens†, Susan E. Andrew\*, Gordon B. Hutchinson\*, Biaoyang Lin\*, Jane Theilmann\*, Rona Graham\*, Martha L. Glaves†, Elizabeth Starr\*, Helen McDonald\*, Jamal Nasir\*, Keith Schappert†, Michael A. Kalchman\*, Lorne A. Clarke\* & Michael R. Hayden\*‡

\* Department of Medical Genetics, University of British Columbia, Vancouver V6T 1Z4, Canada

† Department of Genetics, The Hospital for Sick Children, Toronto M5G 1X8, Canada

**HUNTINGTON'S disease (HD)** is a late-onset autosomal dominant neuropsychiatric disorder presenting in mid-adult life with personality disturbance and involuntary movements, cognitive and affective disturbance, and inexorable progression to death<sup>1</sup>. The underlying genetic defect has been mapped to chromosomal band 4p16.3 (refs 2, 3). Analysis of specific recombination events in some families with HD has further refined the location of the HD defect to a 2.2 megabase DNA interval<sup>4,5</sup>. Using a direct complementary DNA selection strategy we have identified at least seven transcriptional units within the minimal region believed to contain the HD gene. Screening with one of the cDNA clones identified an Alu insertion in genomic DNA from two persons with HD which showed complete cosegregation with the disease in these families but was not found in 1,000 control chromosomes. Two genes including the previously identified  $\alpha$ -adducin gene and another that encodes for a 12-kilobase transcript, map in close proximity to the Alu insertion site. The 12-kilobase transcript should be regarded as a strong candidate for the HD gene.

The distal boundary of the region of interest is defined by a crossover occurring between D4S98 and D4S43 (ref. 4) and the proximal boundary by recombination between D4S10 and D4S125 (ref. 5; Fig. 1). Nonrandom allelic association detected by D4S95 further refines the possible location of the HD gene<sup>6,7</sup>; moreover a second candidate region distal to D4S111 has been designated<sup>8,9</sup>.

We have analysed 250 HD families for recombination events between different markers to define more precisely the minimal candidate region likely to contain the HD gene. In a single family, a recombinational event distal to D4S125 has been identified (Fig. 2). In this family with clearly documented HD in numerous individuals, the recombination event in individual H-6 implies that the mutation causing HD is distal to marker D4S125 thus significantly reducing the candidate region for the HD mutation (Fig. 1).

To identify genes located in HD candidate regions, we have used a direct cDNA selection strategy on three overlapping yeast artificial chromosomes (YACs) which span the region of interest with minimal overlap<sup>10-12</sup> (see legend to Fig. 1 for details). Clones were physically mapped and categorized into transcription units by cross-hybridization to RNA from a variety of tissues and cell lines and to each other. The results for seven GT clones are shown in Fig. 2. Of the clones that were isolated from the 70D11 YAC, one group corresponds to the  $\alpha$ -adducin gene previously identified<sup>12,14</sup>.

Several of the remaining clones contain open reading frames (ORFs). In total, at least seven transcription units with different

transcript sizes could be defined by cDNA clones in the tissues tested. A subset of these clones map to the proximal candidate region between D4S127 and D4S182 which contains D4S95 and also encompasses DNA markers from the core haplotype that is present on about one third of disease chromosomes<sup>15</sup>.

We then tested for rearrangements with those GT clones that map to BIN3 which contains D4S127, D4S95 and D4S182 (see Fig. 1). One GT clone, GT48 detects an insertion of 331 base pairs (bp) in two of 250 HD patients. This rearrangement segregated with HD in both families (Fig. 3a and b) and was seen in genomic DNA digested with various enzymes (Fig. 4c). In one of these families (Fig. 3a) recombination had placed the HD gene distal to D4S125 (Fig. 1b). The rearrangement was not seen in 1,000 control chromosomes. The fact that this rearrangement occurred on the same haplotype in both families and that this haplotype was unique among 140 HD families further suggested a common origin for this rearrangement (Fig. 1b). Both families were of Scottish origin with their ancestors living 50 km apart. The same core chromosomal haplotype extending for about 1 megabase (Mb) including alleles at D4S95 and D4S98 is seen in 2% (14/687) chromosomes of the general population. Normal individuals with chromosomes with this core haplotype do not have this rearrangement.

This rearrangement is localized to a 1.2-kilobase (kb) *HindIII* fragment which contains a portion of GT48 (Fig. 4a). In both families, sequence analysis revealed a 331-bp inserted element which is a member of the Alu family of mobile repetitive elements. With primers flanking the insertion site, the inserted element could be detected using the polymerase chain reaction (PCR; Fig. 4b).

Only a few Alu elements have been described which are transcriptionally active *in vivo* of which the HS Alu subfamily represents the most recently inserted group<sup>16,17</sup>. DNA sequence analysis of this Alu rearrangement (Fig. 4a) shows an insertion by retrotransposition as the Alu element is flanked by a perfect 9-bp duplication of the target sequence, characteristic of insertions of mobile elements into staggered single-stranded nicks at different genomic locations. Furthermore, the sequence corresponding to the insertion site is AT rich, consistent with the hypothesis that Alu elements preferentially integrate into AT-rich regions<sup>18</sup> (Fig. 4a).

Several GT clones allowed the identification of a cDNA for the  $\alpha$ -adducin gene<sup>12</sup>. The 3' untranslated region (UTR) of  $\alpha$ -adducin maps 20-kb telomeric to D4S95<sup>12</sup> (Fig. 4c). An oligonucleotide primer which spans nucleotides 38-58 in the 5' UTR of the  $\alpha$ -adducin gene<sup>14</sup> maps telomeric to the Alu insertion and is located on the same 7.4-kb *EcoRI* fragment as GT24 but does not hybridize to GT24 (Fig. 4c). In addition, a 501-bp RT-PCR product corresponding to nucleotides 38-539 of the  $\alpha$ -adducin cDNA also detected the 7.4-kb *EcoRI* fragment. This places the 5' UTR of the  $\alpha$ -adducin gene close to D4S182, flanking GT24 and indicates that the  $\alpha$ -adducin gene spans at least 80 kb between D4S95 and D4S182 (Fig. 4c).

Because the sequence spanning the Alu insertion showed no identity with the coding region of  $\alpha$ -adducin, the insertion has occurred within an intron, near the 5' end of the  $\alpha$ -adducin gene. No quantitative or size alterations in the  $\alpha$ -adducin gene have previously been detected in RNA prepared from HD heterozygotes and a potential homozygote for HD<sup>13</sup>. It is unlikely, therefore, that the Alu insertion affects the expression of the  $\alpha$ -adducin gene product.

Corresponding transcript(s) for GT48 and two other adjacent clones, GT44 and GT49, were not detected. Northern blot analysis and screening of 10 different cDNA libraries with these cDNA clones did not yield any positive results. Further, sequence analysis of the 1.2-kb *HindIII* fragment containing GT48 did not reveal significant coding potential<sup>19</sup>.

Nevertheless, the presence of a new Alu element might interfere with expression of other genes near the site of insertion. We therefore focused our attention on two other cDNA clones,

† To whom correspondence should be addressed.



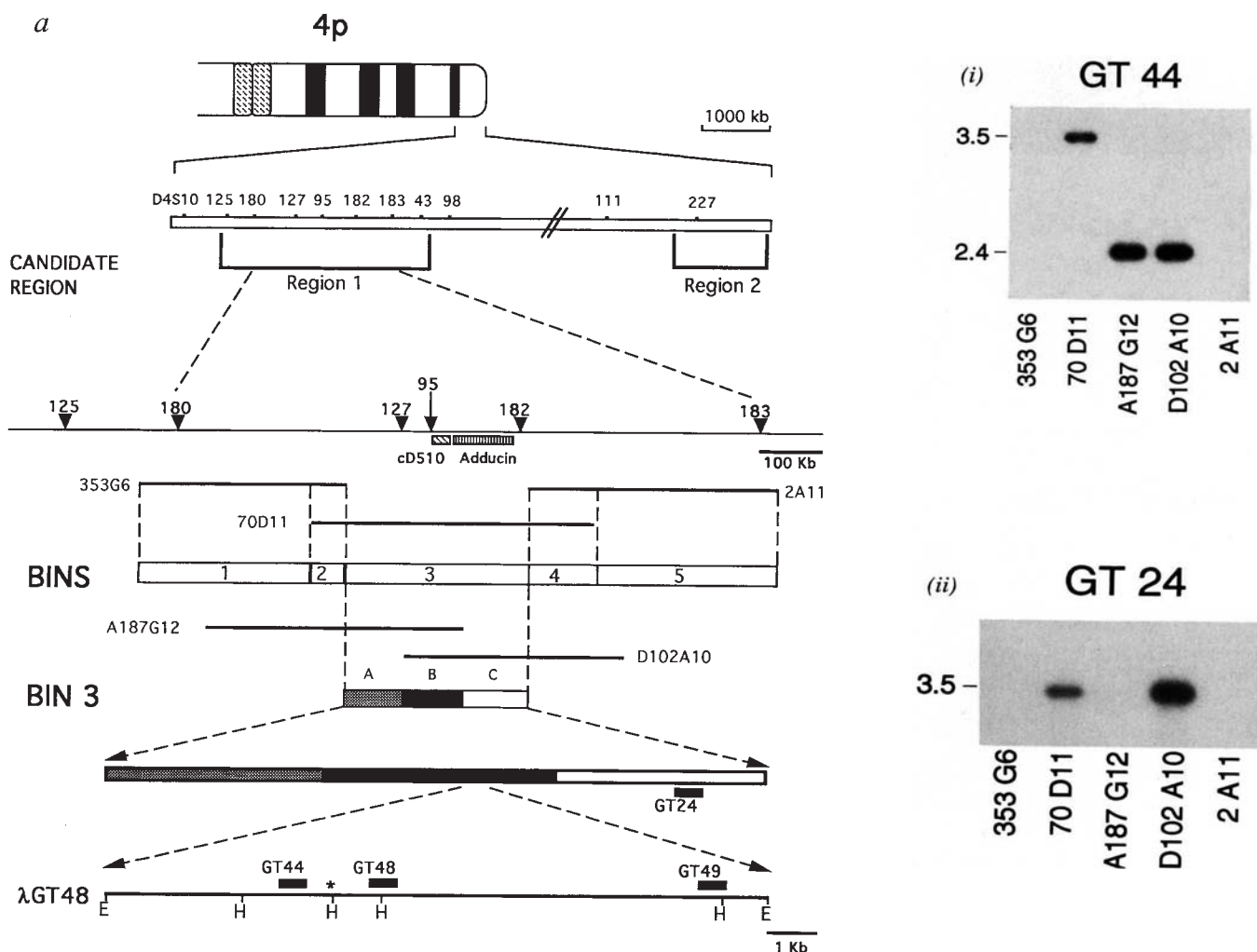
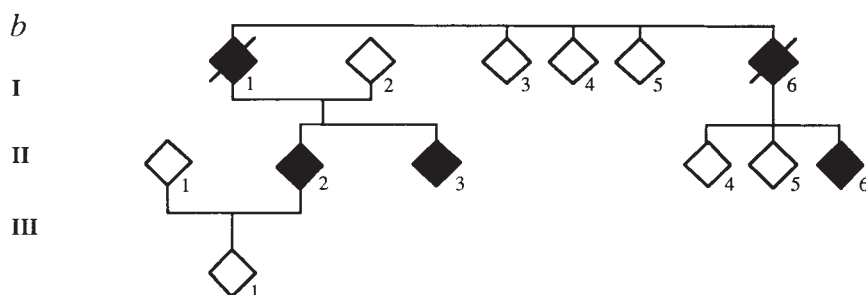


FIG. 1 *a*, Mapping of transcriptional units within the HD candidate region. *b*, Recombination within a HD chromosome refining the proximal boundary of the HD candidate region.

**METHODS.** *a*, DNA markers from 4p16.3 are shown. Purified YACs immobilized on filters were incubated with single or combinations of primary cDNA pools made from four different tissues including fetal brain, frontal cortex, liver and bone marrow. Following two rounds of hybridization, libraries of selected cDNAs were prepared, arrayed and screened for the presence of repetitive sequences as described previously<sup>13</sup>. The origin of the clones was confirmed by hybridization to *Eco*RI-digested DNA from human, human-hamster hybrid containing only human chromosome 4 and the original YAC clones. Each clone was also mapped to a panel of DNAs of overlapping YACs<sup>10,11</sup>. Overlapping regions with YACs 353G6, 70D11 and 2A11 were used to define five separate genomic BINS. YACs A187G12 and D102A10 were used further to refine BIN 3 into three separate compartments: A, B and C. GT clones were mapped by hybridization to digested YAC DNA and assigned to BINS accordingly. GT 44, 48 and 49 mapped to both A187G12 and D102A10, as well as 70D11. All of these clones were contained within a  $\lambda$  phage ( $\lambda$ GT48) isolated using GT 48 as a probe.  $\lambda$ GT48 contains a *Hind*III polymorphism (\*) detected by both GT 44 and GT 48. *a i*, Hybridization pattern of GT 44 to YAC DNAs. *a ii*, GT 24 hybridizes to 70D11 and D102A10 indicating its position within BIN 3C. *b*, The affected haplotype in this family is designated within the boxed region. Haplotypes were determined by analysis of offspring. Recombination between markers D4S125 (YNZ32) and D4S127 in individual H-6 reduced the candidate region of the HD gene. D4S180 was uninformative. For the markers reflecting restriction fragment length polymorphisms (RFLPs). Southern blot analysis was done using previously described methods and probes<sup>24</sup>. Micro-satellite repeats were detected by <sup>32</sup>P-labelled PCR products resolved on a 6% denaturing polyacrylamide gel.



|                |      |      |      |       |       |    |    |      |
|----------------|------|------|------|-------|-------|----|----|------|
| EcoRI (G8)     | BB   | BA   | BB   | BA    | BB    | —  | BB | AB   |
| HindIII (G8)   | DC   | CC   | CB   | BC    | AC    | AC | AA | AA   |
| Bgl I (G8)     | AB   | AB   | AB   | BB    | AA    | AA | AB | AA   |
| PstI (YNZ32)   | 65   | 3/7  | 34   | 4/7   | 85    | 55 | 54 | 1/5  |
| XmnI (D4S180)  | AA   | BB   | BB   | BB    | AA    | AB | AB | BA   |
| PCR - (D4S127) | 22   | 3/5  | 35   | 5/5   | 25    | 53 | 22 | 2/5  |
| AccI (674)     | AB   | AA   | AA   | AA    | AB    | BA | AA | AA   |
| TaqI (674)     | AB   | AA   | AB   | BA    | AB    | BA | BA | BA   |
| MboI (674)     | AB   | BA   | BA   | AA    | AB    | BA | AA | AA   |
| SstI (731)     | BC   | BA   | BA   | AA    | BB    | BB | BA | BA   |
| PstI (252)     | —    | 3/3  | 33   | 3/3   | —     | —  | —  | 1/3  |
| MspI (678)     | AB   | BB   | BB   | BB    | AB    | BA | —  | AB   |
| PstI (157)     | 32   | 2/1  | 22   | 2/1   | 22    | —  | —  | 2/1  |
| PstI (C13B)    | 71   | 2/7  | 22   | 2/7   | 22    | 62 | 66 | 7/7  |
| PCR - (E24CA)  | 11,3 | 3,11 | 3,11 | 11,11 | 11,11 | —  | —  | 7,11 |
| HindIII (2R3)  | 1,2  | 2/2  | 22   | 2/2   | 22    | 22 | —  | 1/2  |

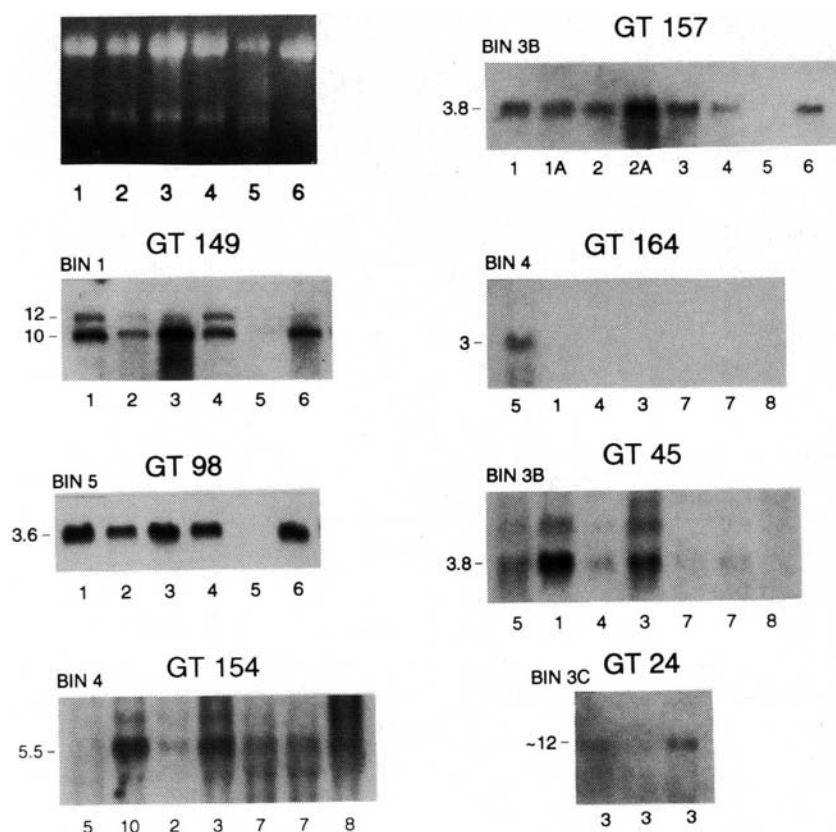


FIG. 2 Northern blot analysis of GT clones. Examples of mRNAs detected with GT clones originating from the candidate HD region are shown. Total RNA from each cell line or tissue was prepared by standard procedures<sup>25</sup>. The lanes represent RNA from the following sources: (1) (10) Caco-2 intestinal cells, (1A) Caco-2 poly (A)<sup>+</sup> RNA, (2) HL60 cells, (2A) HL60 poly (A)<sup>+</sup>, (3) lymphoblasts, (4) fibroblasts, (5) liver, (6) Cos cells, (7) frontal cortex, (8) fetal brain. RNA was separated on 1% agarose gels containing 0.6 M formaldehyde and transferred onto Hybond (Amersham) membranes. The integrity of the RNA is shown by the ethidium bromide-stained gel in the left upper panel. Clones were radio-labelled by random priming and hybridization and washing conditions were as previously described<sup>1,3</sup>. The size of the message detected with each clone is indicated in kilobases and the map location of the GT clone is noted.

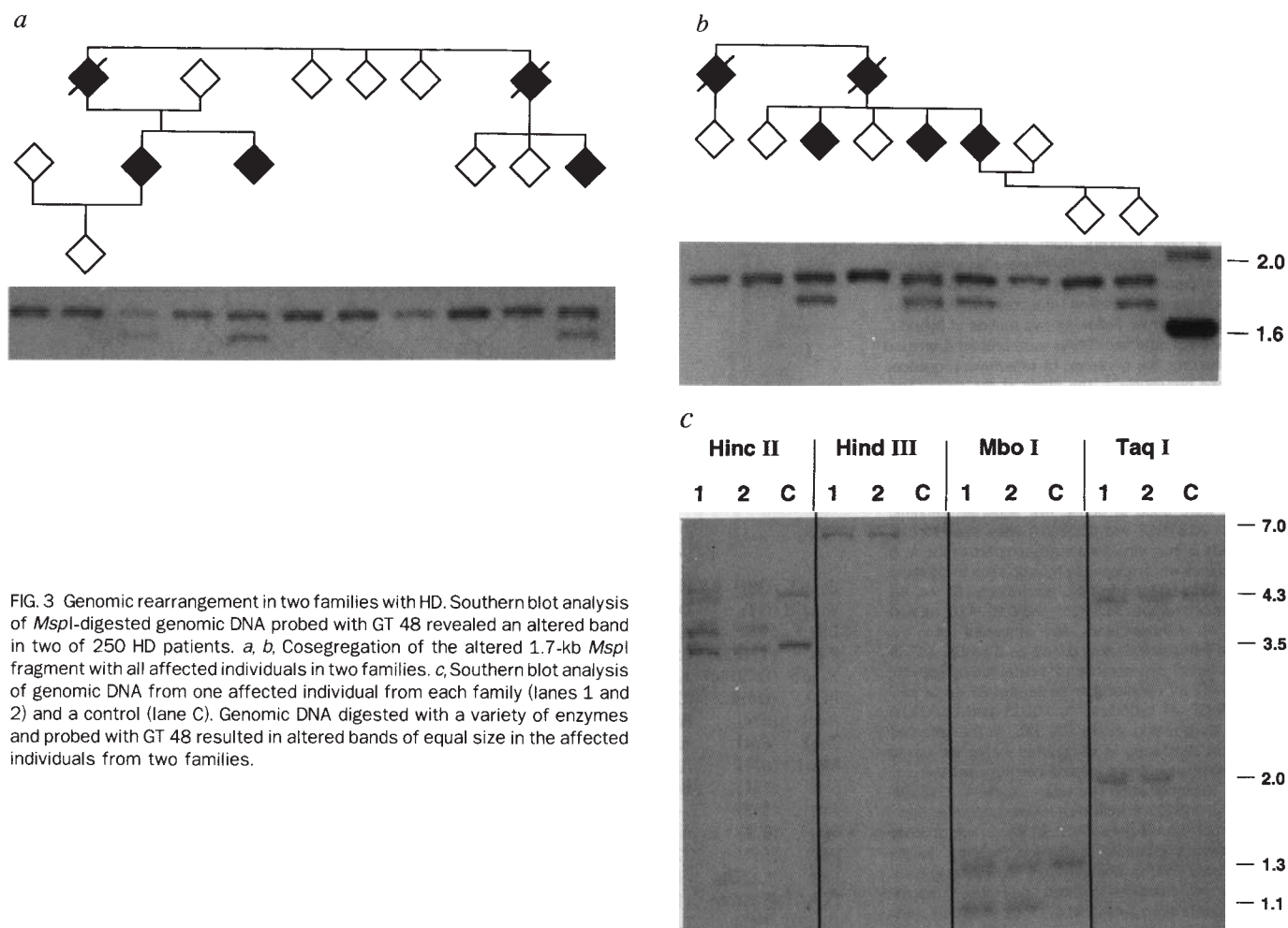


FIG. 3 Genomic rearrangement in two families with HD. Southern blot analysis of *MspI*-digested genomic DNA probed with GT 48 revealed an altered band in two of 250 HD patients. *a, b*, Cosegregation of the altered 1.7-kb *MspI* fragment with all affected individuals in two families. *c*, Southern blot analysis of genomic DNA from one affected individual from each family (lanes 1 and 2) and a control (lane C). Genomic DNA digested with a variety of enzymes and probed with GT 48 resulted in altered bands of equal size in the affected individuals from two families.



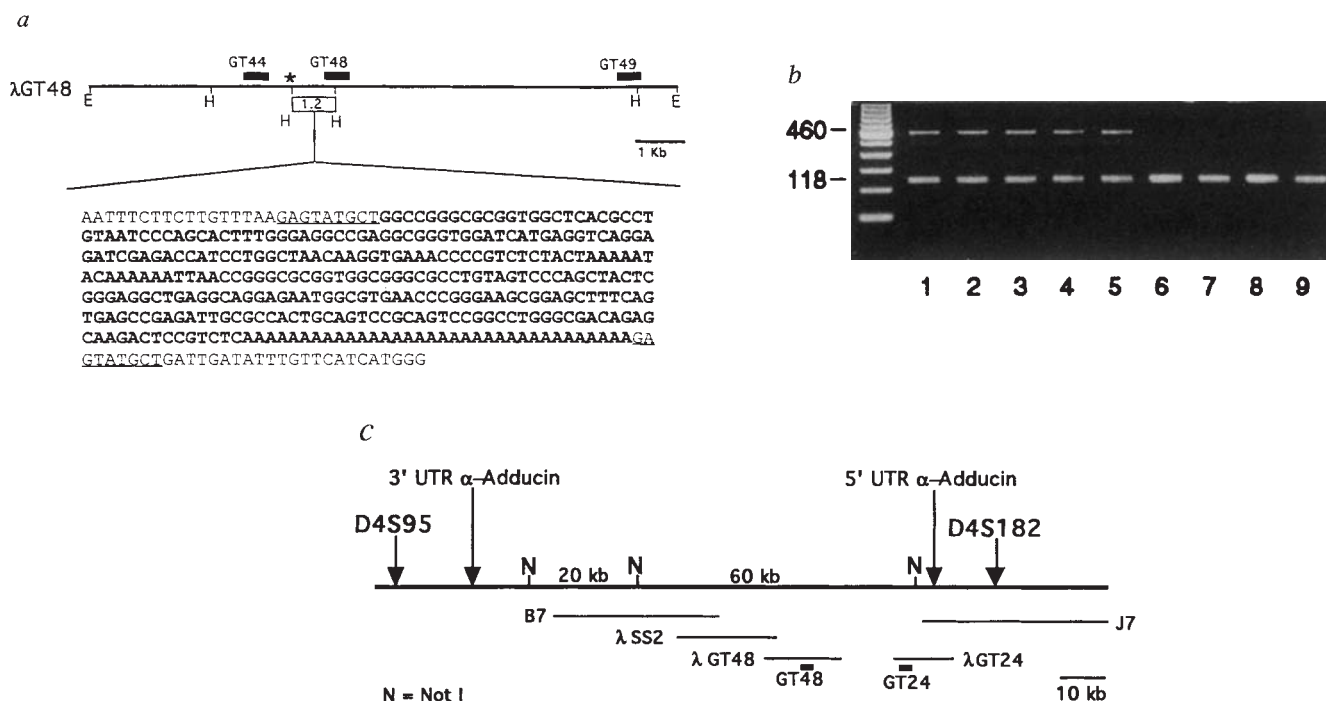


FIG. 4 *a, b*, Alu retrotransposition within the HD candidate region. *c*, Physical map between D4S95 and D4S182. *a, b*, Mapping of the genomic region around GT 48 in controls and the affected individuals localized the rearrangements to the 1.2-kb *Hind*III fragment on λGT48 (boxed). The 1.2-kb *Hind*III fragment was subcloned, sequenced and PCR primers spanning the insertion site were derived. These primers (A, ATGTAATTGTTTCAGGACATGTGGC; B, AAATAACATCCAGAATCTTCAGAT) generated a 118-bp fragment in normal individuals (*b*, lanes 6–9) and in addition a 460-bp product in five affected individuals from both families (*b*, lanes 1–5). The 460-bp PCR product was subcloned (TA cloning, Invitrogen) and sequenced by ABI automated sequen-

cing. The inserted sequence represents a full-length Alu element (bold) and the insertion site is flanked by a 9-bp direct repeat (underlined). *c*, Long-range physical mapping localized GT 24, GT 48 and the α-adducin cDNA clone to the same 60–70-kb *Not*I fragment. Cosmids J7 and B7 were isolated from a chromosome-4-specific library with D4S182 and α-adducin cDNA, respectively, λGT48 and λGT24 were isolated from a λ phage library using their respective GT clones. λGT48 and λSS2 form a contig overlapping with cosmid B7. An oligonucleotide from the 5' UTR of adducin detected the D4S182 cosmid J7 as well as λGT24. By physical mapping GT 48 is about 20 kb from GT24.

GT24 and GT34. Northern blot analysis showed that GT34 detected a 4-kb transcript in a variety of tissues including brain, lymphoblasts and fibroblasts. A 1.7-kb cDNA clone (cD510) was then isolated with a GT34 probe. Sequence analysis of this cDNA clone revealed no homology with sequences in Genbank. The genomic DNA sequence corresponding to cD510 maps distal to D4S95, but centromeric to the 3' UTR of α-adducin and approximately 70 kb from the site of the Alu insertion (Fig. 1), and is therefore an unlikely candidate for the HD gene.

The third clone, GT24, was mapped about 20 kb from GT48 (Fig. 4c). Although GT24 is contained in an intron of the α-adducin gene, it detected a transcript of 12 kb (Figs 1 and 2) in many tissues including frontal cortex, fibroblasts, lymphoblasts, and intestinal cells. Besides some weak identity with the LINE-1 element, this clone also has no homology with any sequence in the data base, but a 69-bp ORF flanked by appropriate splice junctions was noted<sup>20</sup>. Moreover, based on its map position close to the Alu insertion site, the 12-kb transcript is a

candidate gene for HD.

It is not clear how the observed insertion may cause HD. Alu transposition may be involved in the causation of disease by interruption of exonic sequence as for the cholinesterase<sup>21</sup> and factor VIII<sup>22</sup> genes or alternatively it may be located within an intron of a gene where potential splice sites could affect processing of the primary transcript as seen for the insertion in the neurofibromatosis 1 gene<sup>23</sup>.

GT24 located within an intron of the α-adducin gene recognizes a large transcript. It is unlikely that this transcript is located within a single intron of the α-adducin gene which would indicate that these two genes are overlapping, representing a novel genomic organization. Gene mapping of cD510, 70 kb from the Alu insertion, and previous assessment of α-adducin as a candidate provides evidence against these genes being involved in HD. But the transcribed sequence detected by GT24 adjacent to the rearrangement must be further assessed as a candidate gene for HD. □

Received 8 March; accepted 16 March 1993.

- Hayden, M. R. *Huntington's Chorea* (Springer, New York, 1981).
- Gusella, J. F. et al. *Nature* **306**, 234–238 (1983).
- Wang, H. S. et al. *Am. J. hum. Genet.* **3**, 392–396 (1986).
- Snell, R. G. et al. *Am. J. hum. Genet.* **51**, 357–362 (1992).
- Bates, G. P. et al. *Am. J. hum. Genet.* **49**, 7–16 (1991).
- Snell, R. E. et al. *J. med. Genet.* **26**, 673–675 (1989).
- Theilman, J. et al. *J. med. Genet.* **26**, 676–681 (1989).
- Weber, B. et al. *Nature Genet.* **2**, 216–222 (1992).
- Robbins, C. et al. *Am. J. hum. Genet.* **44**, 422–425 (1989).
- Bates, G. T. et al. *Nature Genet.* **1**, 180–187 (1992).
- Zuo, J. et al. *Hum. molec. Genet.* **1**, 149–159 (1992).
- Lovett, M., Kere, J. & Hinton, L. M. *Proc. natn. Acad. Sci. U.S.A.* **88**, 9628–9632 (1991).
- Goldberg, Y. P. et al. *Hum. molec. Genet.* **1**, 669–675 (1992).
- Taylor, S. A. M. et al. *Nature Genet.* **2**, 223–227 (1992).
- MacDonald, M. E. et al. *Nature Genet.* **1**, 99–103 (1992).
- Batzer, M. A. & Deininger, P. C. *Genomics* **9**, 481–487 (1991).

- Materz, A. G. *Nucleic Acids Res.* **18**, 6091–6023 (1990).
- Daniels, G. R. & Deininger, P. C. *Nucleic Acids Res.* **13**, 8939–8954 (1985).
- Überbacher, E. C. & Mural, R. J. *Proc. natn. Acad. Sci. U.S.A.* **88**, 11261–11265 (1991).
- Hutchinson, G. B. & Hayden, M. R. *Nucleic Acids Res.* **20**, 3453–3462 (1992).
- Muratani, K., et al. *Proc. natn. Acad. Sci. U.S.A.* **88**, 11315–11319 (1991).
- Vidaud, D. et al. *Eur. J. hum. Genet.* **1**, 30–36 (1993).
- Wallace, M. R. et al. *Nature* **353**, 864–866 (1991).
- Andrew, S. et al. *Genomics* **13**, 301–311 (1992).
- Chomezynski, P. & Sacchi, N. *Analyt. Biochem.* **162**, 156–159 (1987).

ACKNOWLEDGEMENTS. Y.P.G., J.M.R. and S.E.A. made equal contributions to this project. We thank L.-C. Tsui and R. Worton for comments, G. P. Bates, R. M. Myers and D. R. Cox for YACS, L.-C. Tsui and P. Ray for access to phage libraries, and M. L. Klimek and T. Garber for providing some family DNA. This work is supported by the MRC (Canada), the Canadian Genetic Disease Network and the Huntington Society of Canada. We also acknowledge technical assistance from J. McArthur, V. Lai and Y. Ma, and support from J. Scheuer. Y.P.G. is a Sandoz PDF and M.R.H. is an established investigator for the B.C. Children's Hospital.

PLEASE follow these guidelines so that your manuscript may be handled expeditiously.

*Nature* is an international journal covering all the sciences. Contributors should therefore bear in mind those readers who work in other fields and those for whom English is a second language, and write clearly and simply, avoiding unnecessary technical terminology. Space in the journal is limited, making competition for publication severe. Brevity is highly valued. One printed page of *Nature*, without interruptions, contains about 1,300 words.

Manuscripts are selected for publication according to editorial assessment of their suitability and reports from independent referees. They can be sent to London or Washington and should be addressed to the Editor. Manuscripts may be dealt with in either office, depending on the subject matter, and will where necessary be sent between offices by overnight courier. High priority cannot be given to presubmission enquiries; in urgent cases they can be made in the form of a one-page fax. All manuscripts are acknowledged on receipt but fewer than half are sent for review. Those that are not reviewed are returned as rapidly as possible so that they may be submitted elsewhere without delay. Contributors may suggest reviewers; limited requests for the exclusion of specific reviewers are usually heeded. Manuscripts are usually sent to two or three reviewers, chosen for their expertise rather than their geographical location. Manuscripts accepted for publication are typeset from the London office.

*Nature* requests authors to deposit sequence and crystallographic data in the databases that exist for this purpose, and to mention availability of these data.

Once a manuscript is accepted for publication, contributors will receive proofs in about 4 weeks. *Nature*'s staff will edit manuscripts with a view to brevity and clarity, so contributors should check proofs carefully. Manuscripts are generally published 2–3 weeks after receipt of corrected proofs. *Nature* does not exact page charges. Contributors receive a reprint order form with their proofs; reprint orders are processed after the manuscript is published and payment received.

## Categories of paper

**Review articles** survey recent developments in a field. Most are commissioned but suggestions are welcome in the form of a one-page synopsis addressed to the Reviews Coordinator. Length is negotiable in advance.

**Articles** are research reports whose conclusions are of general interest and which are sufficiently rounded to be a substantial advance in understanding. They should not have more than 3,000 words of text (not including figure legends) or more than six display items and should not occupy more than five pages of *Nature*.

Articles start with a heading of 50–80 words written to advertise their content in general terms, to which editors will pay particular attention. The heading does not usually contain numbers, abbreviations or measurements. The introduction to the study is contained in the first two or three paragraphs of the article, which also briefly summarize its results and implications. Articles have fewer than 50 references and may contain a few short subheadings.

**Letters** are short reports of outstanding novel findings whose implications are general and important enough to be of interest to those outside the field. Letters should have 1,000 or fewer words of text and four or fewer display items. The first paragraph describes, in not more than 150 words and without the use of abbreviations, the background, rationale and chief conclusions of the study for the particular benefit of non-specialist readers. Letters do not have subheadings and should contain fewer than 30 references.

**Commentary articles** deal with issues in, or arising from, research that are also of interest to readers outside research. Some are commissioned but suggestions can be made to the Commentary Editor in the form of a one-page synopsis.

**News and Views articles** inform non-specialist readers about new scientific advances, sometimes in the form of a conference report. Most are commissioned but proposals can be made in advance to the News and Views Editor.

**Scientific Correspondence** is for discussion of topical scientific matters, including those published in *Nature*, and for miscellaneous contributions. Priority is given to letters of fewer than 500 words.

## Preparation of manuscripts

All manuscripts should be typed, double-spaced, on one side of the paper only. An original and four copies are required, each accompanied by artwork. If photographs are included, five sets of originals are required; for line drawings, one set of originals and four good-quality photocopies are acceptable. Reference lists, figure legends and tables should all be on separate sheets, all of which should be double-spaced and numbered. Three copies of relevant manuscripts in press or submitted for publication elsewhere should be included with submitted manuscripts, and clearly marked as such. Five copies of revised and resubmitted manuscripts, labelled with their manuscript numbers are required, together with five copies of a letter detailing the changes made.

**Titles** are brief and simple. Active verbs, numerical values, abbreviations and punctuation are to be avoided. Titles should contain one or two key words for indexing purposes.

**Artwork** should be marked individually and clearly with the author's name and, when known, the manuscript number. Ideally, no figure should be larger than 28 by 22 cm. Figures with several parts are to be avoided and are permitted only if the parts are closely related, either experimentally or logically. Unlettered originals of photographs should be provided. Suggestions for cover illustrations, with captions and labelled with the manuscript number, are welcome. Original artwork is returned when a manuscript cannot be published.

Protein/nucleotide sequences should ideally be in the three-letter and not the single-letter code for amino acids. One column width of *Nature* can accommodate 20 amino acids or 60 base pairs.

**Colour artwork.** A charge of £500 per page is made as a contribution towards the cost of reproducing colour figures. Inability to pay these costs will not prevent publication of essential colour figures if the circumstances are explained. Proofs of colour artwork may be sent to contributors under separate cover from their galley proofs.

**Figure legends** should not exceed 300 words and ideally should be shorter. The figure is described first, then, briefly, the method. Reference to a method published elsewhere is preferable to a full description. Methods are not described in the text.

**References** are numbered sequentially as they appear in the text, followed by those in tables and finally by those in figure legends. Only papers published or in the press are numbered and included in the reference list. All other forms of reference should be cited in the text as a personal communication, manuscript submitted or in preparation. Text is not included in reference lists. References are abbreviated according to the *World List of Scientific Periodicals* (Butterworths, London, 1963–65). The first and last page numbers are included; reference to books should include publisher, place and date.

**Abbreviations**, symbols, units and greek letters should be identified the first time they are used. Acronyms should be avoided whenever possible and, if used, defined. Footnotes are not used in the text.

**Acknowledgements** are brief and appear after the reference list; grant and contribution numbers are not allowed.

**Supplementary information** is material relevant to Articles or Letters which cannot, for lack of space, be published in full, but which is available from *Nature* on request.

**Submission.** Manuscripts can be sent to the Editor at 4 Little Essex Street, London WC2R 3LF, UK or at 1234 National Press Building, Washington, DC 20045, USA. Manuscripts or proofs sent by air courier to London should be declared as 'manuscripts' and 'value \$5' to prevent the imposition of import duty and value-added tax.



# Eye on the marketplace

**Somatic cell hybrid panels, a new immunodeficient mouse model, reagents for dopamine and glutamate receptor research, cell proliferation assays, and a 750-MHz nuclear magnetic resonance spectrometer are featured.**

ONCOR announces the release of new **somatic cell hybrid panels** to assist in the assignment of human DNA probes to specific human chromosomes (*Reader Service No. 100*). Each panel contains DNA obtained from the best available monochromosomal human-rodent hybrid cell lines restricted with either *Pst*I, *Bam*HI or *Msp*I. Rodent and human controls, molecular weight size markers, and an individual hybrid to every human chromosome (1 through 22, X and Y) are included on a single membrane. Tested for chromosome content and purity, Oncor says that these panels are easy to use and interpret. The somatic cell hybrid panels are transferred onto Oncor Sure Blot membrane — a positively charged nylon membrane that



**Somatic cell hybrid panels from Oncor.**

permits stripping and reprobing a minimum of five times. An ethidium bromide picture of the gel electrophoresis and an instruction manual are provided. Oncor has also introduced a **large fragment probe labelling kit** that is designed to label efficiently cosmids and yeast artificial chromosomes for fluorescence *in situ* hybridization. Large fragments are nick translated to the ideal size for performing cytogenetic mapping studies including primary localization and ordering of probes. The kit includes cold nucleotides, diluting and stop solutions, and enzyme mixture to label up to 20 µg of DNA. Oncor says that each reaction requires only 75 min to complete and can be used with biotin- or digoxigenin-labelled nucleotides. Useful information on optimizing the probe size is contained in the instruction booklet. Labelled nucleotide must be purchased separately.

Becton Dickinson offers a comprehensive line of **monoclonal antibodies for studying leukocyte-endothelial interactions** (*Reader Service No. 101*). The CAMFolio line-up includes  $\beta$ 2 integrins, E-selectin (anti-ELAM-1), P-selectin (anti-GMP-140), anti-



**Becton Dickinson's portfolio of products for leukocyte-endothelial research.**

ICAM-1, anti-integrin  $\alpha$ 4 (VLA-4), L-selectin (anti-LECAM-1), anti-VCAM-1 and anti-sialyl-L<sup>x</sup>. Additional CAMFolio monoclonal antibodies for cell adhesion research have recently been released, including anti-integrin  $\alpha$ 2, anti-integrin  $\beta$ 3, anti-integrin  $\beta$ 4, CD45RO, mouse IgG1 control, mouse IgG2a control and mouse IgG2b control. CAMFolio monoclonal antibodies are specific for human cell adhesion molecules and are manufactured in a high-purity, ready-to-use formulation for multiple applications such as adhesion blockade, ELISA and immunoprecipitation. Becton Dickinson says that they are purified to a minimum specification of 95 per cent purity as determined by sodium dodecyl sulphate-polyacrylamide gel electrophoresis, and are provided at 1.0 mg ml<sup>-1</sup> concentration with no added carrier proteins. Clone-specific references are provided, as are specificity/reactivity information and a bibliography.

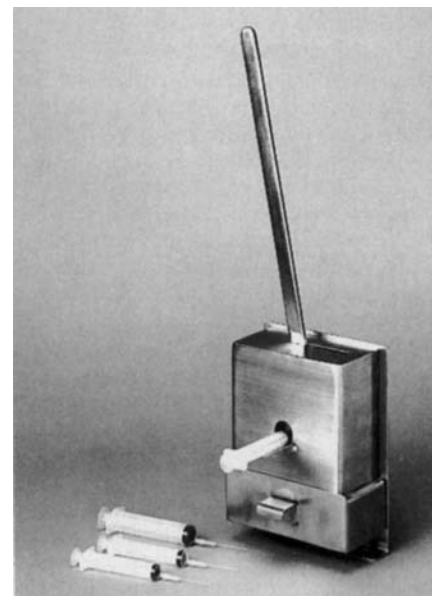
Taconic has added a new **immunodeficient mouse**, an outbred lcr-SCID, to its line of immunodeficient animal models (*Reader Service No. 102*). The lcr-SCID is an outbred mouse that is suitable for research involving tumour xenografts and ascites production. In addition to its lack of T and B cells, the lcr-SCID has a larger body size, superior reproductive rate and lower cost than SCID mice maintained on inbred backgrounds (such as the C.B-17-SCID). lcr-SCID breeder stock was received at



**More mice models from Taconic.**

Taconic from the Fox Chase Cancer Center in early 1992. Germfree caesarian derivations were performed at Taconic and the derived pups were associated with the Altered Schaedler Flora. Genetic monitoring tests on the derived pups confirm their outbred status and homozygosity for the SCID gene.

Balcan Engineering has produced a new machine for the quick, safe and easy **destruction of all sizes of used hypodermic needles** (*Reader Service No. 103*). Called the Balcan needle destructor, the device is



**Balcan's Destructor is designed to reduce the risk of needlestick injuries.**

manually operated and intended to be mounted on any surface or ward trolley. To use, the hypodermic syringe is inserted into the side of the device and the operating handle is pulled towards the operator. This causes several specially designed blades to chop the needle into many short pieces of less than 5 mm in length. Both the plastic hub of the needle and the nib of the syringe are cut off. This also ensures that the syringe cannot be reused. The fragments of the needle fall into a small plastic pot underneath which, when filled, can be capped ready for proper disposal. Once the needle has been destroyed, the syringe is removed and disposed of as clinical waste plastic. The destructor device is fabricated throughout of stainless steel so that it can be easily cleaned. Balcan says that destroying needles in this way overcomes many of the problems (such as needlestick injuries, the risk of theft of contaminated needles and their use by unauthorized

personnel) associated with the disposal of used needles because they no longer remain intact.

The AxioDoc A digital image archives from Carl Zeiss links the microscope, stored images and the connected data in a **digital archiving system** that is based on a PC with a 486 processor and the Windows 3.1 user interface (*Reader Service No. 104*). The archival system consists of an additional card for image acquisition, the MaRS (magnification recognition system) system for Zeiss microscopes with ICS (infinity colour-corrected system) optics and image filing software. The archives can contain many different 'boxes', depending on the number of users or subjects. This means that images can be inseparably interlinked with external data, such as numbers of specimens. Furthermore, it is possible to perform simple measurements of the images and to store the results, such as lengths and angles.

### ■ Receptor research

Research Biochemicals has licensed the exclusive rights to produce and market the **human D<sub>3</sub> dopamine receptor** (*Reader Service No. 105*). The company hopes that because high levels of D<sub>3</sub> receptors are found in the limbic areas of the brain, the D<sub>3</sub> receptor has become a therapeutic target in the search for new antipsychotic agents. The human D<sub>3</sub> dopamine receptor is expressed in a murine fibroblast cell line and developed through recombinant cell culture technology. RBI says that the commercial availability of the D<sub>3</sub> receptor will allow researchers access to this receptor without having to isolate or clone the receptor.

Glutamate is the major excitatory neurotransmitter in the mammalian central nervous system. At least four classes of receptors mediate the action of glutamate in the CNS. The major class of receptor, the AMPA receptor, is comprised of four distinct subunits. Chemicon's new **rabbit polyclonal antibodies to the glutamate receptors (GluR)** are now available (*Reader Service No. 106*). These include anti-GluR 1, anti-GluR 2 and 3 and anti-GluR 4. The antibodies are supplied affinity purified and can be used for immunocytochemistry on para-formaldehyde/glutaraldehyde-fixed tissue. They may also be used for western blot and immunoprecipitation studies.

### ■ Immunoassays

Advanced Magnetics has introduced a single range competitive **enzyme immunoassay kit** for the quantitative determination of adenosine 3':5'-cyclic monophosphate (cAMP) levels in biological fluids (*Reader Service No. 107*). cAMP, an important signal transduction compound, has been shown to play a role in the cardiovascular and nervous system, in immune mechanisms, cell growth and metabolism. Intracellular cAMP measurement in tissues and cell cultures is



AMI's cAMP enzyme immunoassay kit.

yielding further insight into the physiology and pathology of many disease states. Advanced Magnetics says that the total assay time is 6 h and requires no acetylation. The EIA is optimized with a sensitivity of 20 fmol ml<sup>-1</sup>. Each kit is supplied with precision-coated wells and contains two-plate ELISA-containing reagents sufficient for 192 tests, including a standard curve.

AMAC's **interleukin-6 assay** is a one-step 'sandwich' enzyme immunoassay (EIA) that utilizes monoclonal antibodies (*Reader Service No. 108*). The sample and conjugate are added to the monoclonal-coated microtitre plate and incubated for 2 h at room temperature on a shaker. After washing the microtitre plate, a chromogenic substrate is added and incubated at room temperature for 30 min on a shaker. The plate can then be read on a spectrophotometer between 405 nm and 414 nm. AMAC says that this IL-6 EIA has a sensitivity of 0.5 pg ml<sup>-1</sup>. The kit is supplied in a 96-test kit format containing all the reagents necessary to perform the assay.

### ■ Assorted reagents

Over 15,000 compounds can be found in the latest edition of the Fluka Chemie **catalogue** for chemists, biochemists and analysts (*Reader Service No. 109*). Users are provided with detailed and useful chemical and biochemical data. New safety information such as the Swiss Toxicity Classification, the German Water Hazard Classification, in addition to handling/storage recommenda-



Shelf stockers from Fluka.

tions have been added. Identification is facilitated by the CAS Registry Numbers and structural data. The new Beilstein Registry Number is a key for online research within the Beilstein database. Literature references and citations (Merck etcetera) further enhance the handbook character of this catalogue. The catalogue is also available on diskette.

Fermentech has taken nine years to fine tune its continuous fermentation process using naturally occurring *Staphylococcus aureus* to produce a pure form of **Protein A** (*Reader Service No. 110*). The product's specifications make it suitable for clinical research or therapeutic use, while it should be attractive as a monoclonal cleansing agent because it is free from *Escherichia coli* and other contaminants encountered with the genetically engineered product. Supplied as a ready soluble lyophilized powder or immobilized onto Sepharose gel, Fermentech says its Protein A has a greater than 98 per cent purity as measured by HPLC. In addition, there are no detectable changes in purity or activity even after storage for 3 years at -20 °C. Normal temperature storage conditions for the material range between 4 and 8 °C. Each consignment is supplied with a certificate of analysis, identification information and a detailed UV spectrum.

**Protein kinase C** plays a central role in the transduction of growth signals and in the regulation of cellular functions. The enzyme, which is available from Boehringer Mannheim, is purified from rat brain and contains a mixture of the isoforms  $\alpha$ ,  $\beta$ 1,  $\beta$ II and  $\gamma$  (*Reader Service No. 111*). Boehringer Mannheim says that PKC is greater than 90 per cent pure (by SDS-PAGE) and has a specific activity of greater than 1 unit per milligram (histone as the substrate). The preparation is available as a solution without the addition of protein stabilizers, and is free of proteases. Protein kinase C from rat brain is suitable for the characterization and detection of potential substrates and phosphorylation sites, for screening new activators and inhibitors, for the production of phosphoproteins and phosphopeptides (phosphatase substrates), and for mechanistic investigations. Related products that are also available include protein kinase C inhibitory peptide (19-31), protein kinase C substrate [Ac-MBP (4-14)] (which is specific for protein kinase C and resistant towards protein phosphatases), histone H1 and staurosporin. Protein kinase A (PKA, cAMP-dependent protein kinase) phosphorylates a broad spectrum of target proteins with preference for the sequence motif Arg-Arg-Xxx-Ser/Thr-Xxx. Whereas the PKA holoenzyme consists of a tetrameric structure (R<sub>2</sub>C<sub>2</sub>), which has to be activated by cAMP, the preparation represents the catalytic subunit which is not regulated and exhibits the full activity. The enzyme is highly purified from bovine brain (greater than 95 per cent by



SDS-PAGE) and possesses a specific activity of greater than 6 units per milligram (kemptide as the substrate). It is supplied as a solution filled under nitrogen for optimal convenience. Boehringer Mannheim says that protease activity is not detectable. The enzyme is suitable for elucidating the mechanisms of neural and hormonal signal transduction. Protein kinase A substrate (kemptide) is also available.

Two new nonradioactive products for the **measurement of cell proliferation** have been introduced by Amersham (*Reader Service No. 112*). The new cell proliferation



**Assays for measuring cell proliferation.**

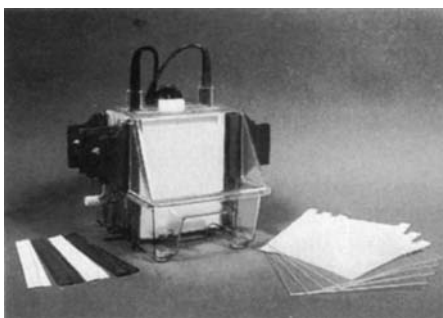
kit is an immunocytochemical system for monitoring cell proliferation in tissue specimens derived from tissue culture and whole animal experiments. Amersham says that some of the benefits of the new kit include time savings (proliferation can be detected in less than 4 h), high sensitivity that is up to autoradiographic standards, a choice of *in vitro* and *in vivo* studies, enhanced signal through lowered competition by endogenous thymidine, and greater information per sample through multiple labelling experiments. For the quantification of cell proliferation in cultured cells, the company has introduced the cell proliferation assay. Additional stated benefits include reproducible high performance through quality-controlled matched reagents and, being nonradioactive, there are no problems with the disposal of waste scintillant.

### ■ Electrophoresis

If you are in need of high-resolution RNA gels for northern analysis, Bioplex has developed an **electrophoresis system**, called STAG, which is based on the surface-tension technique (*Reader Service No. 113*). Bioplex says that the system allows for the optimal resolution of different RNAs in thin agarose gels and guarantees sharp bands after transferring to hybridization filters. Fourteen different samples can be loaded onto one gel with the maximal amount of 15 µg RNA

per lane. The package includes an electrophoresis unit, glass plates, a levelling table, hybridization filter, a buffer recirculation pump and all parts and information necessary to cast and run surface-tension gels for northern blot analysis.

The SE260 slab gel unit from Hoefer Scientific, with its new gel format of 8-cm wide × 8.5-cm long, combines the ease and simplicity of the Mighty Small Vertical Slab Gel format with the greater separation of longer gels (*Reader Service No. 114*). The



**Hoefer expands its Mighty Small line with the new SE260 slab gel unit.**

chamber can be used for both acrylamide and agarose gels used in conventional, pulsed-field or isoelectric focusing applications. It runs two gels at a time, with up to 15 samples each, in about 45 min, Hoefer says. The moulded upper buffer chamber integrates a serpentine heat exchanger and clamp-free tubing connectors. The deep lower buffer chamber accommodates precast gels of up to 10 cm long. The company also makes gels casters for casting up to ten gels for the SE260 at once.

### ■ Separation systems

Bps separations announces the launch of its Productiv IMP (inverted matrix purification) products, extending the range of Productiv 10-ml cartridges to include products for the **medium- and large-scale purification of proteins** (*Reader Service No. 115*). The product range is based on the use of a monolithic sponge or matrix to support the separation chemistries. Liquid flows through large interconnecting pores in the continuous matrix as opposed to around the beads of conventional chromatographic supports. Bps says that the principal benefit of this configuration is the ability to load the crudest extracts, from cell lysates to whole broths, directly onto the IMP columns, without the need for sample pretreatment. Additional benefits include high flow rates and low back pressures, which, combined with the tolerance to cell debris, make this product suitable as an initial concentration step for target proteins present in high volume, low concentrations or, alternatively, for the rapid removal of contaminating proteins. The additions to the range are supplied as column prepacks of 50-mm depth for commonly available columns up to diameters of

#### ADVERTISEMENTS

#### SINGER MSM SYSTEM

for YEAST GENETICS:-

- Ascus Dissection
- Pedigree Analysis
- Cell & Zygote Isolation

**SINGER**  
INSTRUMENTS

"A responsibility to Science"

Call or Fax now for further information  
& demonstration video

Tel: +44 984 40226 Fax: +44 984 41166

Singer Instrument Co Ltd, Roadwater, Watchet, Somerset, England, TA23 0QL

**Reader Service No.13**

#### Synthetic Oligonucleotides

Rapid Quality Worldwide Service

**£2.50 per base, no setup!**

option of H.P.L.C. purification  
biotinylation, fluorescent labelling  
RNA Synthesis

phosphothioates, many other modifications

#### OSWEL DNA SERVICE

Department of Chemistry, University of Edinburgh  
Kings Buildings, West Mains Road,  
EDINBURGH EH9 3JJ

Tel: 031 650 4792 Fax: 031 662 4054

**Reader Service No.15**

#### Highest Quality Oligonucleotides

Synthesised, purified and shipped in 3 days

**£2.75 per base**

(All inclusive)

0.2 µM scale (approx. 500 µg)

- Antisense (completely or partially sulphurised)
- Extensive range of labelling and modifications
- Full design and analysis facilities
- Gene cassettes, bulk order and contract discounts

Molecular Medicine Unit

King's College School of Medicine & Dentistry

Tel. 071 - 326 3126 FAX 071 - 733 3877

**Teflon<sup>\*</sup>  
Coated  
and**

**Bare Wire from  
MEDWIRE**

Medwire provides  
the highest quality  
platinum, platinum alloy, silver and stainless  
steel wire in both single and multistrand  
configurations.

For prompt delivery or additional  
information, please contact:

**MEDWIRE**, an affiliate of Sigmund Cohn Corp.  
121 So. Columbus Ave. Mount Vernon, NY 10553  
Tel: (914) 664-5300 Fax: (914) 664-5377

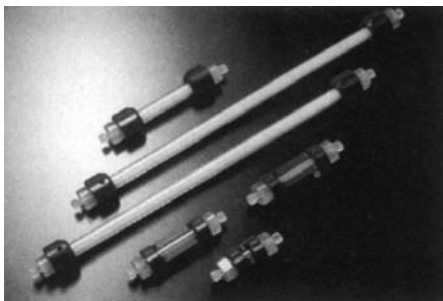
\*Reg TM of E I DuPont & Co

**Reader Service No.25**

## PRODUCT REVIEW

250 mm. These prepacks can be added to a column to any desired depth without detriment to flow or media compression, bps says.

TosoHaas features offers TSK-GEL PW polymeric resins available in glass columns for use with FPLC and HPLC systems (*Reader Service No. 116*). With a particle size of 10 mm and a 1,000-Å pore size, the TSK-GEL PW resins offer high resolution in addition to good flow properties for protein and other biomolecular separations. The columns have been developed for size



TSK-GEL glass columns for FPLC.

exclusion, ion exchange, hydrophobic interaction and affinity chromatography. A wide range of functional chemistries are available, including DEAE, SP, CM, phenyl, butyl, heparin, chelate and tresyl. TSK-GEL glass columns are available in three sizes from 5-mm to 20-mm internal diameter and with endfittings ready to attach to any FPLC systems. Endfittings for HPLC are available separately.

### ■ Detection systems

Varian has developed a 750-MHz nuclear magnetic resonance spectrometer (NMR) that should be useful for determining structures of biomolecules such as proteins, nucleic acids (DNA and RNA) and

system's experimental flexibility and to make upgrades easier.

TopCount is the new two-in-one detection system from Packard that performs both **luminescence and radioisotope counting** in one instrument (*Reader Service No. 118*). This dual-purpose microplate counter can measure glow and enhanced flash luminescence on the same instrument, with no modifications, along with  $^3\text{H}$ ,  $^{125}\text{I}$  and any radioisotope that can be counted by scintillation analysers. Using patented reflective optics and an advanced electronics design, TopCount can be used to analyse luminescence assays such as reporter genes, DNA capture, and immunoassays using black microplates to prevent optical crosstalk and plate phosphorescence. Radioassays such as receptor binding, adherent cells, and in-plate immunoassays can be measured in white microplates for maximum efficiency. To increase throughput, two wells of the microplate are counted simultaneously, a 20-plate sample changer allows untended operation and a barcode reader keeps track of plate identification and protocol selection. The isothermal counting chamber ensures that proper counting assay conditions are maintained.

The new Optima 3000 inductively coupled plasma optical emission spectrometer (ICP-OES) from Perkin-Elmer offers a new detector designed specifically for plasma emission spectroscopy and a patented optical system that enables users to measure the spectral background and each analyte line simultaneously (*Reader Service No. 119*). This allows users to measure 60 elements in less than one minute at multiple wavelengths, without sacrificing precision or sensitivity, the company says. In addition, the Optima 3000, which occupies only one square metre of floor space, includes over 5,000 emission lines, providing the flexibility to select alternate interference-free lines for better results. The spectrometer consists of an Echelle-based polychromator with a segmented-array, charge-coupled detector, a 40-MHz free-running RF generator with true power control and temperature-controlled plasma pneumatics in a compact, free-standing package. The system provides two doors into the sample compartment for removal and replacement of the quick-change torch module. It is controlled from a 486-based computer that automates all instrument functions, including plasma operation, control of the autosampler and data acquisition. The software stores relevant operating parameters and provides post-analysis data processing of stored spectra. □

*These notes are compiled by Diane Gershon from information provided by the manufacturers. For more information, fill in the reader service card bound inside the journal.*

#### ADVERTISEMENTS

**\$3.50/Base**

**Custom DNA  
Synthesis**

No Set-up Fee,  
Shipped in 24 hours.

1-800-788-4362  
1-800-786-4362 Fax

**KEYSTONE**  
Laboratories, Inc.

Reader Service No.33



**Severn  
Biotech Ltd.**

#### CUSTOM SERVICES

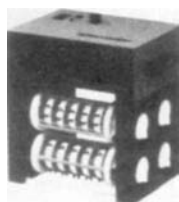
- **OLIGONUCLEOTIDE SYNTHESIS**  
High quality PROBES, PRIMERS, ANTISENSE OLIGOS!  
Fast-Efficient Service offering loads of modifications,  
BIOTIN, FLUORESCENCE, 5' & 3' labelling,  
Phosphorothioates Affinity Support/Oligos, etc.
- **MULTIPLE PEPTIDE SYNTHESIS**  
Epitope Mapping & Screening! We now supply 5mg  
quantities of high purity custom peptides up to 15  
residues in length, using Fmoc chemistry. With D or L  
amino acids, these are supplied lyophilised in 4ml  
vials, cleaved from solid phase & ready for use.

Tel: 0562 825286 Fax: 0562 825284

Unit 23, Heathbrook Enterprise Centre  
Warwick Road, Kidderminster DY10 1HY

Reader Service No.26

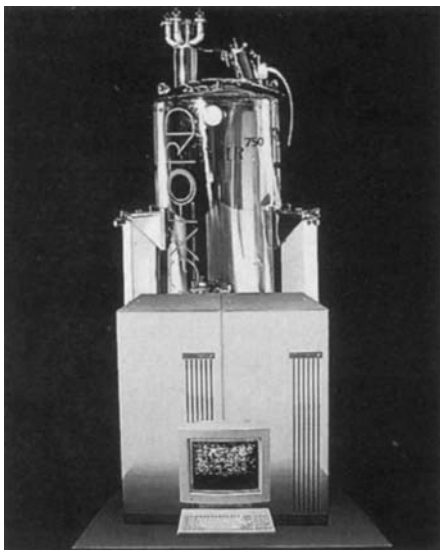
#### **DIABORM®** EQUILIBRIUM DIALYSIS SYSTEM



The gentlest method to quantify the binding parameters and thermodynamic data between ligands and biopolymers. Results come closest to the actual values. Sample volume from 0.2 ml to 5.0 ml. Up to 20 experiments in parallel.

**DIABORM**  
Geräte  
P.O. Box 65076,  
D-8000 Munich 65,  
FRG.

Reader Service No.45



Varian's 750-MHz NMR spectrometer.

carbohydrates (*Reader Service No. 117*). Since 1986, NMR analysis has been performed on spectrometers up to 600 MHz, but Varian says that the results are often difficult or impossible to interpret because of overlapping features in the spectra. The increased dispersion that is offered by Oxford Instruments' 750-MHz magnet in the UNITYplus NMR system is designed to decrease the overlap of peaks in the spectra, allowing researchers to determine the structures of larger molecules. In addition, the increased sensitivity of the 750-MHz system permits the study of smaller sample quantities. The UNITYplus is a modular system, which is intended to increase the